Electronic Properties of Next-Generation Semiconductors

Matthew Z. Mayers

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2018

© 2018
Matthew Z. Mayers
All rights reserved

ABSTRACT

Electronic Properties of Next-Generation Semiconductors

Matthew Z. Mayers

The need for efficient, cheap, and durable semiconductors for photovoltaic and optoelectronic applications has spurred a number of dramatic recent developments in semiconductor quantum physics. Aided by advanced synthetic and characterization techniques, the development of high-quality, nano-structured, tunable materials has resulted in the observation of many novel phenomena. The goal of this thesis is to develop and apply methods in theoretical condensed matter science to the study of these promising materials.

In Chapter 1 I explore methylammonium lead iodide ($MAPbI_3$), a paradigmatic hybrid organic-inorganic perovskite system. To explain charge carrier dynamics in this material, I develop a microscopic tight-binding model. The average band structure is calculated and the magnitude of the temperature-dependent band gap opening and Urbach energy is quantified. The charge carrier mobility is calculated within a linear response formalism and its temperature dependence is characterized. Overall, the fully *ab initio* model is found to explain several non-trivial experimental phenomena while making minimal assumptions concerning the nature of the electron-phonon coupling and the character of the nuclear motion in these materials.

In Chapters 2 and 3, I turn to the subject of atomically-thin transition metal dichalcogenides. I improve upon past variational calculations of exciton and trion binding energies in these materials by applying diffusion Monte Carlo to *exactly* calculate exciton, trion, and biexciton binding energies within an effective few-body Hamiltonian. Carriers are assumed to experience two-body interactions of the Keldysh type that have been parameterized previously from electronic structure calculations. The structures of the exact ground state wavefunctions are calculated and compared to those of the previous variational trial wavefunctions. Next, I calculate the doping dependence of the rate of exciton and trion elastic scattering with free electrons within first-order time-dependent perturbation theory. The calculation provides the first theoretical estimate of the intrinsic trion

linewidth in these materials.

Finally, in Chapter 4, I study variants of the GW approximation to the one-particle Green's function for calculating correlation energies and spectral weights for the three-dimensional homogeneous electron gas. By relating the cumulant generating function to the improper GW self-energy, I develop a new cumulant-based GW approximation. The approach is compared to existing methods first via solution of a simple linearly-coupled electron phonon model and later through application to the electron gas problem.

TABLE OF CONTENTS

List of Figures iv

List of Tables vii

Chapter 1: Carrier Dynamics and Absorption Features in Hybrid Organic-Inorganic Perovskites 1

1.1 Introduction 1

1.2 Tight-binding Hamiltonian 3

1.3 Results: Bandgap Fluctuations 7

1.4 Quantum Dynamics 9

1.5 Results: Electron and Hole Mobilities 11

1.6 Conclusion 15

1.7 Appendix: Initial Conditions 17

1.7.1 Full thermal trace 17

1.7.2 Decomposition of full thermal trace 18

1.8 Appendix: Matrix Element Spectral Densities and non-Gaussianity 19

1.8.1 Effective electron-phonon couplings and spectral coupling functions 19

1.8.2 Non-Gaussian parameter 21

Chapter 2: Binding Properties of Small Carrier Complexes in Two-dimensional Transition Metal Dichalcogenides 22

2.1 Introduction . . . 22

2.2 Screening in two-dimensional transition metal dichalcogenides . . . 24

2.3 Variational Monte Carlo . . . 25

2.4 Diffusion Monte Carlo . . . 26

2.5 Results . . . 33

2.6 Conclusion . . . 40

Chapter 3: Doping-Dependent Exciton and Trion Linewidths . . . 41

3.1 Introduction . . . 41

3.2 Methods . . . 42

3.2.1 Fermi's Golden Rule . . . 42

3.2.2 Exciton-electron scattering matrix elements . . . 42

3.2.3 Trion-electron scattering matrix elements . . . 44

3.2.4 Material parameters for MoS_2 . . . 46

3.3 Results . . . 46

3.4 Conclusions . . . 48

Chapter 4: A New GW Cumulant Approximation . . . 49

4.1 Introduction . . . 49

4.2 The electron gas (jellium) model . . . 52

4.3 The Green's function . . . 54

4.4 Results . . . 61

4.5 Conclusion . . . 67

4.6 Appendix: Spurious Sharp Quasiparticle Resonances at $k > k_F$ . . . 68

4.7 Appendix: Derivation of Electron-Phonon Green's Function 69

Bibliography 74

LIST OF FIGURES

1.1 Band structure for the tight-binding model (a) without and (b) with spin-orbit coupling using average matrix elements $\overline{\epsilon_{\text{Pb s}}} = -2.5$ eV, $\overline{\epsilon_{\text{I p}}} = 2.5$ eV, $\overline{\epsilon_{\text{Pb p}}} = 6.0$ eV, $\overline{t_{\text{sp}\,\sigma}} = 1.0$ eV, $\overline{t_{\text{pp}\,\sigma}} = 1.8$ eV, $\overline{t_{\text{pp}\,\pi}} = 0.4$ eV, $\gamma_{\text{Pb}} = 0.56$ eV, $\gamma_{\text{I}} = 0.31$ eV. 4

1.2 Illustrations of the tight-binding Hamiltonian parameterization based on atomic coordinates. (a) Time series for the fluctuations in a typical diagonal (blue, "on-site energy") and off-diagonal (black, "hopping") matrix element. The fluctuations in the hopping parameters are approximately a factor of 2 larger. (b) Best exponential fit (blue curve) to a scatter plot of the sp hopping matrix element vs. Pb-I bond distance. The data were taken from a DFT calculation on snapshots of the crystal structure obtained from MD calculations. (c) Typical time series for an sp hopping matrix element. The insets show snapshots from the underlying MD trajectory at the marked times with the relevant atoms enlarged. The pink atoms represent iodine, gold atoms represent lead, and MA cations are show in blue and white. (d) Plots of the matrix element $\epsilon_i^\alpha(\boldsymbol{R}(t))$ vs. t calculated using DFT (black) and classical electrostatics (colored) for a Pb s (bottom) and Pb p (top) orbital. 5

1.3 Characterization of (a) the temperature-dependent bandgap opening and (b) the temperature-dependent Urbach energy E_{U} compared in each case to experimental work. In (a), AHC+TE represents the standard Allen-Heine-Cardona estimate of vibrationally-induced bandgap change. 8

1.4 A summary of the periodic boundary conditions used in our model. Quantum dynamics are performed within the full supercell A, using matrix elements determined from a molecular dynamics trajectory performed in supercell B that are periodically replicated within supercell A. In calculating the on-site energy for an orbital of a particular atom (black dot in the diagram), an Ewald summation is performed using only atomic coordinates within supercell C centered on that atom. 10

1.5 Carrier mobilities are calculated using the diffusion constant for spreading of the charge carrier. Heat maps for a single charge carrier density within a cross-section of the 3D system are shown at $t = 5$ fs in panel (a) and at $t = 50$ fs in panel (b). The simulation was initialized with a conduction band Wannier function, which is localized within about 1 nm. The spread of the carrier for this simulation is plotted vs. time in panel (c), with the data points corresponding to panels (a) and (b) highlighted. Using data averaged over multiple initial conditions, the temperature dependence of the mobility, plotted in panel (d), was found to most closely fit $T^{-2.11}$. Experiments suggest a temperature dependence of the form T^{-x} with $1.4 \lesssim x \lesssim 2.8$ [35, 36, 37, 38]. 12

1.6 The function $\langle\Psi(t)|\widehat{\delta r}^{\,2}|\Psi(t)\rangle$, plotted for a single trajectory and initial electronic wavefunction. The wavefunction was propagated using different versions of $\hat{H}_{\text{TB}}$. . 13

1.7 The charge carrier mobility, calculated using Eq. 1.9 for electrons and calculated using Eq. 1.12 for both electrons and holes. The data was fitted to a curve of the form aT^{-b}. The best estimate of the mobility of electrons at $T = 300$, calculated using the full thermal trace, is about 135 cm^2/Vs. 14

1.8 The real and imaginary parts of $MSD(t)$, calculated using a partial trace in Eq. 1.9. 17

1.9 Spectral densities calculated at $T = 300$ K using the time-correlation function of the parameterized tight binding Hamiltonian matrix elements for (a) hopping amplitudes and (b) onsite matrix elements. In panel (b), the Pb s and Pb p on-site matrix elements yield identical spectral densities because the parameterized tight binding model assumes the fluctuations in those terms to be the same. 19

1.10 Calculated NGP for (a) on-site energies and (b) hopping amplitudes. In panel (a), the Pb s and Pb p on-site matrix elements yield identical non-Gaussian parameters because the parameterized tight binding model assumes the fluctuations in those terms to be the same. 21

2.1 A schematic depiction of the DMC imaginary time propagation process for the calculation of the ground state trion energy. Each square corresponds to a single walker at a point in imaginary time. Electrons (holes) are depicted with filled (unfilled) circles. As only the relative electron and hole coordinates are significant, the hole is always depicted in the same position. 29

2.2 (a) Radial probability distributions for the distance r_{eh} of an exciton in MoS_2. (b) Radial probability distributions for the distances r_{eh} and r_{ee} of a negative trion in MoS_2. (c) Radial probability distributions for the distances r_{eh}, r_{ee}, and r_{hh} of a biexciton in WSe_2. For the DMC calculation, the curves for r_{ee} and r_{hh} coincide because the electron and hole effective masses are taken to be equal. 35

2.3 (a) Radial probability distributions for the distances r_{eh} and r_{ee} in the trion (X^-) and biexciton (XX) using a Keldysh form for the inter-carrier potential. (c) The same r_{ee} distributions as in panel (a), but the relocated repulsive weight is shaded. (b),(d) The same as in panels (a),(c) but using a Coulombic form for the inter-carrier potential. 38

3.1 The doping-dependent linewidth broadening due to elastic electron-exciton scattering at $T = 50$ K and $T = 290$ K. 47

3.2 The doping-dependent linewidth broadening due to elastic electron-trion scattering at $T = 50$ K and $T = 290$ K. 47

4.1 Feynman diagrams for the zeroth- and first-order terms in the Green's function expansion. 56

4.2 Definitions of the (a) proper self-energy and (b) proper polarization insertion. . . . 58

4.3 The (a) self-energy and (b) polarization insertion used in the non-self-consistent GW (or G_0W_0) approximation. 58

4.4 A comparison of three distinct GW schemes. (a,c) The spectral weight for the 3D electron gas with $r_s = 4.0$. (b,d) The absolute value of the imaginary part of the proper self-energy, solved using Dyson's equation, corresponding to the spectral weights in (a) and (c). 63

4.5 Band structure for various GW schemes at $r_s = 4$. Note that the G_0W_0 IRC contains multiple satellites, similar to the G_0W_0 PRC case, but with a spacing that mimics the incorrect position of the G_0W_0 plasmaron peak. 64

4.6 The $k-$dependent occupation number $n(k)$ for all three approximation schemes at $r_s = 4.0$, plotted against the "exact" QMC result [145, 146]. 64

4.7 Correlation energies of the 3D electron gas per particle ϵ_{corr}/N for various GW schemes and r_s values compared with the "exact" QMC result. Exact values are reported in Table 4.1. 66

4.8 Band structure for various GW schemes at $r_s = 4$ and $k = 3k_F$. The sharp quasiparticle resonance predicted by G_0W_0 IRC is broadened by the bIRC and rIRC schemes. 68

4.9 Spectral weight for the Green's function of the electron-phonon model within the Einstein phonon model for (a) $g = 0.2$ and (b) $g = 5$. Each vertical line represents a Dirac delta function with weight corresponding to the height of the line. 73

LIST OF TABLES

1.1 Best-fit parameters for the parameterization of the hopping matrix elements $t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t))$ of the form $ae^{-r/b} + c$. 6

1.2 Partial charges used in the classical electrostatic calculation of the on-site matrix elements $\epsilon_j^{\alpha}(\boldsymbol{R}(t))$ [29]. The average on-site energy was calculated using DFT to be -2.5 eV for the Pb s orbital, 2.5 eV for the I p orbitals, and 6.0 eV for the Pb p orbitals. The unit charge is $e = 1.6 \times 10^{-19}$ C. 6

1.3 Temperature dependence of the average bandgap and spread of bandgaps. Bandgaps were calculated using instantaneous nuclear configurations obtained from MD simulations performed on an $8 \times 8 \times 8$ supercell at the various temperatures. The data are converged with respect to supercell size. 7

1.4 Calculated average mobilities for dynamics run with different versions of $\hat{H}_{\text{TB}}$. . . . 15

2.1 Estimated exciton (X) and trion (X^-) binding energies for different members of the 2D TMDC class of materials. Where two numbers are reported, the number on the left is the most current estimate. The statistical uncertainty in the DMC data is on the order of 0.1–0.3 meV. The column labeled "variational" refers to results based on the Keldysh form, taken from Ref. [70]. 33

2.2 Estimated biexciton (XX) binding energies for different members of the 2D TMDC class of materials. The statistical uncertainty in the DMC data is on the order of 0.1–0.3 meV. 33

2.3 Comparison of trion binding energies for several potential forms in units of meV obtained from DMC, except for the column labeled "variational"; the latter results are based on the Keldysh form and taken from Ref. [70]. Binding energies in the "variational 1" column are with respect to the variational exciton binding energies, whereas those in the "variational 2" column are with respect to the exact DMC exciton binding energies. For the DMC Keldysh and pure logarithmic potentials, the uncertainty is of the order of 0.1–0.3 meV. For the pure Coulombic potential, the uncertainty is of the order of 10 meV. 35

2.4 Comparison of biexciton binding energies for several potential forms in units of meV obtained from DMC, except for the column labeled "variational"; the latter results are based on the Keldysh form and taken from Ref. [57] for the WSe_2 biexciton. Binding energies in the "variational 1" column are with respect to the variational exciton binding energies, whereas those in the "variational 2" column are with respect to the exact DMC exciton binding energies. For the DMC Keldysh and pure logarithmic potentials, the uncertainty is of the order of 0.1–0.3 meV. For the pure Coulombic potential, the uncertainty is of the order of 10 meV. 36

4.1 Correlation energies of the 3D electron gas per particle ϵ_{corr}/N in Hartrees for various GW schemes and r_s values [142]. QMC values are obtained from Vosco et al's parameterization [147] of Ceperley and Alder's fixed-node diffusion Monte Carlo data [128]. 66

4.2 Correlation energies of the 3D electron gas per particle ϵ_{corr}/N in Hartrees for the broadened IRC (bIRC) and the renormalized IRC (rIRC) schemes. 68

Acknowledgements

Pursuing a PhD in Chemical Physics means that I've spent much of the past five years inside my own head, willing myself to crack complex derivations and to find pesky bugs in my code that were causing so much trouble. While I learned a lot during those moments, there's no doubt that I grew most significantly through my interactions with my friends and colleagues both at Columbia and elsewhere, without whom I would not be who I am today.

First, I was lucky to be able to work under the most relatable advisor I could hope for, David Reichman. Professionally, Dave is a walking encyclopedia of scientific literature who is also able to convey the personalities and stories of many seminal figures in the field. Personally, he is such a caring and understanding person that when my friends would complain about their advisors, I could not participate in the conversation for lack of substantial complaints of my own.

I am also indebted to the other members of my committee, who each oversaw aspects of my scientific development. Bruce Berne and Angelo Cacciuto were amazing teachers of introductory and advanced statistical mechanics and encouraged me to broaden my research horizons when I could. Mark Hybertsen showed me around Brookhaven National Lab and helped me at the very beginning of my degree to navigate the numerical and conceptual difficulties of many-body perturbation theory. Andy Millis hosted stimulating group meetings, and I appreciate him participating on my committee. My collaborators, including Tim Berkelbach, Andrey Chaves, David Egger, Misha Glazov, Archana Raja, and Alexey Chernikov, were a pleasure to work with and taught me much more about the systems I studied than I could learn from reading literature alone.

The Columbia staff, including the Chemistry department front offices, were extremely helpful with scheduling, reimbursements, and other tasks that I normally have difficulty with. I feel especially indebted to Dr. Persaud, a true master of her trade, who helped me navigate some of my most difficult times here.

Next is Soumyo, my labmate, roommate, and friend who is the kindest soul I know. We've experienced every step of this program together, and I will remember particularly our late night

conversations to the tune of Coldplay. I also thank the rest of the former and current Reichman lab for their advice, help, and good company, including Roel, Yevgeny, Andres, Liesbeth, James, Ian, James, and Benedikt.

My research went through a number of lulls during which time my motivation and curiosity could have become stifled. Somewhat fortuitously, I came across Ryan, a fellow dreamer and innovator, who encouraged and inspired me to explore my passion for virtual reality more deeply and seriously. Our exciting work together kindled the creative flame in me that my scientific research couldn't during my middle years here.

To all of my friends from many different walks of life, thank you. You introduced me to music that brightened my outlook. We played volleyball, softball, soccer, frisbee, Super Mario, and even the marble game until I couldn't play any longer. You encouraged me to grow personally through deeper and more honest introspection.

Finally, I am grateful to have been able to pursue my PhD in a location close to my family. My parents have been a constant source of encouragement and motivation, and I thank them for endowing me with intellectual curiosity and a compulsive attention to detail. My siblings have been patient with my difficulty in relating to their condition even while they understood mine so well. I hope they can both find a peace that allows them to live to the fullest.

For my teachers,

who not only taught

but also joined me in discovery

CHAPTER 1
CARRIER DYNAMICS AND ABSORPTION FEATURES IN HYBRID ORGANIC-INORGANIC PEROVSKITES

1.1 Introduction

Hybrid organic-inorganic lead halide perovskites have emerged in the last decade as a promising alternative to traditional silicon-based semiconductors. Whereas silicon-based photovoltaic devices require expensive fabrication processes to generate crystal structures with minimal defects, the solution-based synthesis of perovskites is cheap and facile, allowing for micron-sized crystals to be grown directly from a solution comprised of abundant materials [1, 2, 3, 4]. In spite of these advantages, perovskite materials are able to suppress many of the efficiency losses that plague organic and dye-sensitized solar cells [2]. Indeed, conversion efficiencies of perovskite photovoltaic devices have increased dramatically from 3.8% in 2009 to over 20% in recent years following straightforward alterations in the synthetic process [3, 4]. In addition, high luminescence yields and facile visible spectral tunability via variation of the material's halide composition make perovskites promising candidates for light-emitting diode and lasing applications [5, 6, 7, 8, 9].

A number of intriguing qualities related to their soft and dynamic structure make perovskites quite different from other crystalline inorganic semiconductors and serve to highlight their uniqueness. First, the material undergoes two first-order phase transitions: a low-temperature transition from an orthorhombic phase to a tetragonal phase near 160 K and a phase change near room temperature from the tetragonal phase to a cubic phase that results in a large decrease in resistivity in the high-temperature material [10, 11]. The observation of octahedral tilting even in single crystals has been linked to the material's interesting phase behavior and to properties of charge carrier dynamics [12]. In addition to these unusual structural properties, perovskites possess a rich collection of strongly anharmonic acoustic and optical phonons [13, 14, 15]. Halide shuttling between

lead atoms occurs on the timescale of about half a picosecond while cation rotation (in the case of a methylammonium lead-halide perovskite) and reorganization lead to dynamic reshaping of the cation cage on slightly longer timescales [16, 11, 15].

These soft, anharmonic vibrational motions are suspected to strongly influence charge carrier dynamics. First, the local polar fluctuations in both Cs and MA perovskites are thought to lengthen the charge carrier lifetimes [13]. Octahedral tilting has also been suggested as a mechanism underlying long carrier lifetimes [12]. On the other hand, the large-amplitude motion and strong electronic coupling to phonon modes is known to be the dominant factor in limiting charge carrier mobilities via the temperature dependence of the mobilities, although there is disagreement over whether coupling to acoustic or optical phonons is more important [15, 17]. In recent years, many studies have proposed the mechanism of large polaron formation to explain the limited mobilities. Theoretical studies parameterized using material properties of perovskites suggest an upper limit of about $100-200$ cm^2/Vs for the electron mobility [18, 15], which is in line with recent experimental observations in single crystals. These Feynman polaron models are, however, inconsistent with the observed temperature-dependence of the charge mobility[18, 15]. Paradoxically, experimental observations of an exceptionally sharp onset of the absorption edge suggest a highly-ordered microstructure [19], an observation which appears to be at odds with the large phonon displacements known to exist in these materials.

In this work, we propose a microscopic tight-binding model for charge carrier dynamics in $MAPbI_3$. The model makes minimal assumptions concerning the harmonicity and the type (acoustic or optical) of the phonon modes that couple to the electronic degrees of freedom. On the other hand, the phonon modes are treated completely classically, thus precluding a large polaron description of the charge carrier dynamics. In this respect, our goal is not to adjudicate the viability of the polaron picture but rather to raise the possibility that many of the interesting phenomena highlighted above can be explained in the absence of polaron formation. The tight-binding model is parameterized using atomic coordinates via density functional electronic structure calculations in Sec. 1.2. We calculate the band structure for our model and quantify the temperature dependence

of the band gap opening and magnitude of the Urbach energy. We then calculate the electron and hole mobilities within linear response theory, paying close attention to the effects of the initial condition upon the outcome. The magnitudes of the calculated mobilities as well as their temperature dependence is found to agree quite well with experiment.

1.2 Tight-binding Hamiltonian

Our tight-binding Hamiltonian employs 13 atomic orbitals per unit cell: an s and three p orbitals on the Pb atom and 3 p orbitals on each I atom [20]. The methylammonium (MA) cation is treated as electronically inert. Each atomic orbital is comprised of two spin states so that the total basis size per unit cell is 26. The system Hamiltonian can be written as a sum of electronic and nuclear components as

$$\begin{aligned} \hat{H}(t) &= \sum_{i,\alpha} \epsilon_i^{\alpha}(\boldsymbol{R}(t)) c_i^{\alpha\dagger} c_i^{\alpha} + \sum_{\langle ij\rangle,\alpha} t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t)) c_i^{\alpha\dagger} c_j^{\alpha} + \hat{H}_{\text{SO}} + H_{\text{nuc}} \\ &= \hat{H}_{\text{TB}} + H_{\text{nuc}}, \end{aligned} \tag{1.1}$$

where $\boldsymbol{R}(t)$ represents all atomic coordinates at time t, the indices i, j refer to atomic orbitals, the index α refers to the orbital spin, and $\langle ij\rangle$ restricts the sum to pairs of orbitals on neighboring Pb and I atoms. More specifically, for each Pb atom, the model includes one sp σ bonds, one pp σ bonds, and two pp π bonds to each of the six neighboring I atoms.

The first term of Eq. 1.1 represents the on-site energy. A linear Taylor expansion of such terms will produce polaronic (Holstein- or Fröhlich-like) couplings if the nuclear motion is quantized [21]. The second term represents the kinetic energy or hopping term. A linear Taylor expansion of this term will produce Peierls or off-diagonal coupling if the nuclear motion is quantized [21]. H_{nuc} represents the lattice energy. We assume the lattice fluctuates in an anharmonic but classical manner. Due to the significant effect of spin-orbit coupling upon the electronic structure of hybrid perovskites [22, 23], it is included in our model via $\hat{H}_{\text{SO}}$, defined by the following

relations:

$$\begin{aligned}
\left\langle p_x^{\uparrow}\middle|\hat{H}_{\mathrm{SO}}\middle|p_y^{\uparrow}\right\rangle &= -\left\langle p_x^{\downarrow}\middle|\hat{H}_{\mathrm{SO}}\middle|p_y^{\downarrow}\right\rangle = i\gamma \\
\left\langle p_x^{\downarrow}\middle|\hat{H}_{\mathrm{SO}}\middle|p_z^{\uparrow}\right\rangle &= -\left\langle p_x^{\uparrow}\middle|\hat{H}_{\mathrm{SO}}\middle|p_z^{\downarrow}\right\rangle = \gamma \\
\left\langle p_y^{\uparrow}\middle|\hat{H}_{\mathrm{SO}}\middle|p_z^{\downarrow}\right\rangle &= \left\langle p_y^{\downarrow}\middle|\hat{H}_{\mathrm{SO}}\middle|p_z^{\uparrow}\right\rangle = -i\gamma,
\end{aligned} \tag{1.2}$$

where $\gamma = 0.56$ eV for Pb, $\gamma = 0.31$ eV for I, and $\uparrow$ ($\downarrow$) refers to the spin-up (spin-down) state.

To calculate tight binding parameters from first principles, the electronic wavefunctions were first determined using density functional calculations with spin-orbit coupling on a $2 \times 2 \times 2$ supercell using snapshots from an MD trajectory as instantaneous nuclear structures. The DFT calculations were performed using the QUANTUM-ESPRESSO package with the PBE generalized gradient approximation exchange-correlation functional [24]. Norm-conserving, designed non-local pseudopotentials were generated with the OPIUM package [25, 26]. A plane-wave cutoff energy of 50 Ry was sufficient to converge the total energy with k-point sampling on a 4×4×4 grid. The tight binding Hamiltonian was built by projecting the computed wavefunctions onto the atomic orbital basis functions obtained using the Wannier90 package [27, 28]. Instead of solving for the maximally-localized Wannier functions, the tight binding Hamiltonian matrix elements were read off from the first step of the Wannierization process, which is the initial projection of

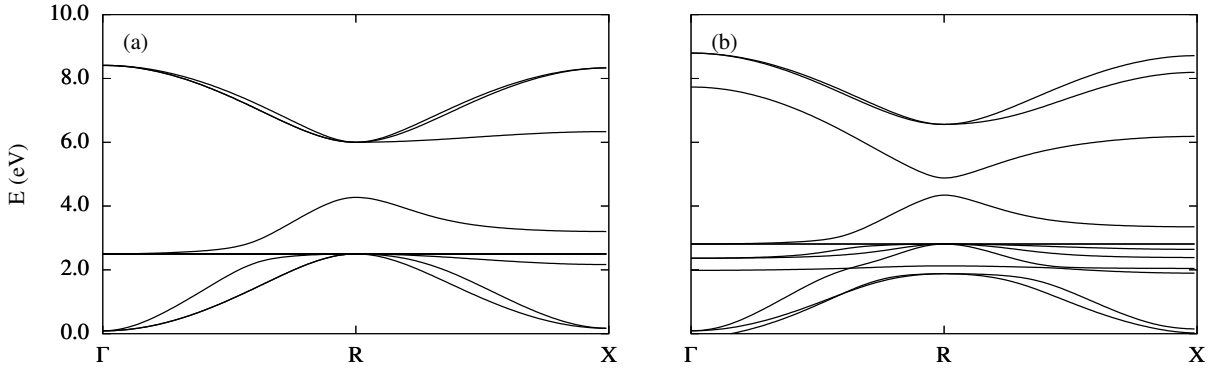

Figure 1.1: Band structure for the tight-binding model (a) without and (b) with spin-orbit coupling using average matrix elements $\overline{\epsilon_{\mathrm{Pb\,s}}} = -2.5$ eV, $\overline{\epsilon_{\mathrm{I\,p}}} = 2.5$ eV, $\overline{\epsilon_{\mathrm{Pb\,p}}} = 6.0$ eV, $\overline{t_{\mathrm{sp}\,\sigma}} = 1.0$ eV, $\overline{t_{\mathrm{pp}\,\sigma}} = 1.8$ eV, $\overline{t_{\mathrm{pp}\,\pi}} = 0.4$ eV, $\gamma_{\mathrm{Pb}} = 0.56$ eV, $\gamma_{\mathrm{I}} = 0.31$ eV.

wavefunctions on the atomic orbital basis.

The time-averaged quantities $\overline{\epsilon_i^\alpha(\boldsymbol{R}(t))}$ and $\overline{t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t))}$ were determined and used to calculate static band structures for the tight-binding model with and without spin-orbit coupling, which are presented in Fig. 1.1. (Note that as discussed in Sec. 1.3, these parameters vary weakly as a function of temperature. For all dynamics calculations, we assume the time-averaged tight binding parameters to be temperature-independent, taking their values at $T = 300$ K as quoted in Fig. 1.1.) Most notably, the bandgap for the model with spin-orbit coupling (0.53 eV) is about $1/3$ that of the model without spin-orbit coupling (1.73 eV) [22]. The experimental bandgap is about 1.5 eV [22].

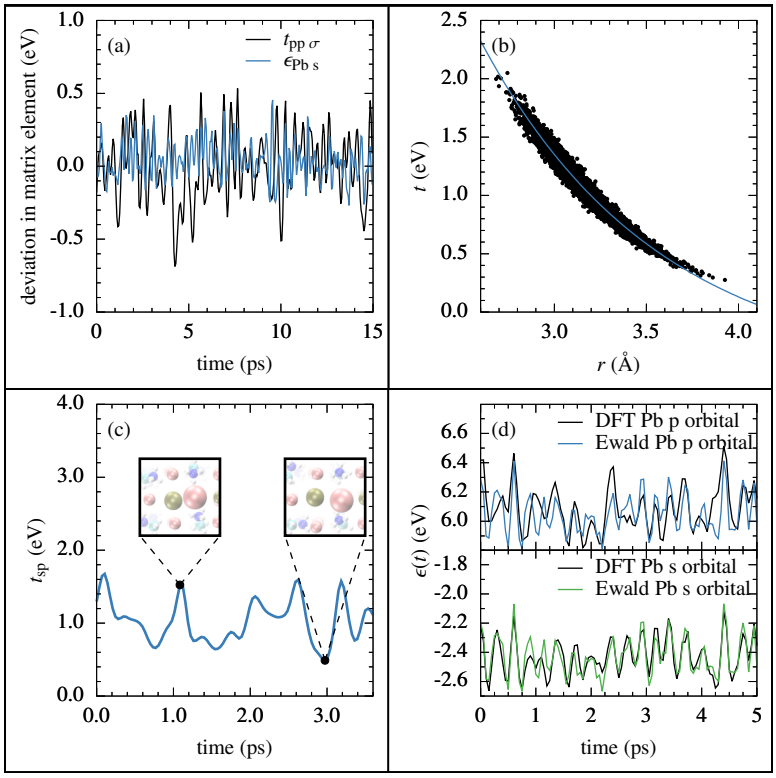

Figure 1.2: Illustrations of the tight-binding Hamiltonian parameterization based on atomic coordinates. (a) Time series for the fluctuations in a typical diagonal (blue, "on-site energy") and off-diagonal (black, "hopping") matrix element. The fluctuations in the hopping parameters are approximately a factor of 2 larger. (b) Best exponential fit (blue curve) to a scatter plot of the sp hopping matrix element vs. Pb-I bond distance. The data were taken from a DFT calculation on snapshots of the crystal structure obtained from MD calculations. (c) Typical time series for an sp hopping matrix element. The insets show snapshots from the underlying MD trajectory at the marked times with the relevant atoms enlarged. The pink atoms represent iodine, gold atoms represent lead, and MA cations are show in blue and white. (d) Plots of the matrix element $\epsilon_i^\alpha(\boldsymbol{R}(t))$ vs. t calculated using DFT (black) and classical electrostatics (colored) for a Pb s (bottom) and Pb p (top) orbital.

hopping type	a (eV)	b ($\mathring{A}$)	c (eV)
sp σ	42.49	0.964	-0.54
pp σ	21.61	3.070	-5.88
pp π	11.66	1.030	-0.16

Table 1.1: Best-fit parameters for the parameterization of the hopping matrix elements $t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t))$ of the form $ae^{-r/b} + c$.

atom	partial charge q/e
Pb	2.03
I	-1.13
C	-1.1
H($-$C)	0.54
N	0.771
H($-$N)	0.023

Table 1.2: Partial charges used in the classical electrostatic calculation of the on-site matrix elements $\epsilon_j^{\alpha}(\boldsymbol{R}(t))$ [29]. The average on-site energy was calculated using DFT to be -2.5 eV for the Pb s orbital, 2.5 eV for the I p orbitals, and 6.0 eV for the Pb p orbitals. The unit charge is $e = 1.6 \times 10^{-19}$ C.

In Fig. 1.2b, we plot the hopping matrix element $t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t))$ vs. $r_{ij}(t) = |\boldsymbol{R}_i(t) - \boldsymbol{R}_j(t)|$, where $\boldsymbol{R}_i(t), \boldsymbol{R}_j(t)$ refer to the coordinates at time t of the atoms containing orbitals i and j. We determined that $t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t))$ could be reasonably parameterized by the function $t_{ij}^{\alpha\alpha}(t) = ae^{-r_{ij}(t)/b} + c$, which takes a single variable $r_{ij}(t)$ and three parameters. The best-fit curve is plotted over the scatter plot data corresponding to the sp hopping element Fig. 1.2b and the best-fit parameters for each type of hopping are presented in Table 1.1.

A typical time series for the sp hopping matrix element is plotted in Fig. 1.2c. Insets show snashots of the crystal structure at the depicted times, with the relevant Pb-I pair enlarged in gold and pink, respectively. Greater orbital overlap corresponds to a higher hopping matrix element and vice-versa. The fluctuations in the largest hopping matrix element (the pp σ overlap) are significantly larger than the fluctuations in the on-site energies, as shown in Fig. 1.2a.

We found that the fluctuations in the on-site energies $\epsilon_i^{\alpha}(\boldsymbol{R}(t))$ could be reasonably approximated by the electrostatic potential at the position $\boldsymbol{R}_i(t)$ due to all surrounding atomic charges (including those on the inert MA cations). We plot in Fig. 1.2d the time series for the value $\epsilon_i^{\alpha}(\boldsymbol{R}(t))$ computed using the DFT and projection procedure outlined above along with the classical electro-

static version of the same matrix element. The partial charges used in the Ewald summation were taken to be those used in the MD force field [29] and are reported in Table 1.2.

1.3 Results: Bandgap Fluctuations

Given the parameterization of $\hat{H}_{\text{TB}} = \hat{H} - H_{\text{nuc}}$ described in Sec. 1.2, it should be noted that the nuclear-driven fluctuations of the tight-binding parameters give rise to a temperature-dependent distribution of bandgaps. As discussed below, the temperature dependence of the mean of this distribution reproduces the experimentally observable phenomenon of bandgap opening with increasing temperature. The width of the distribution may be used to quantify the experimentally observable Urbach energy, E_{U}, which characterizes the sharpness of the absorption edge.

Disordered structures of an $8 \times 8 \times 8$ supercell were generated using classical molecular dynamics performed within the canonical ensemble [29]. For each structure, all Bloch states were calculated explicitly at the Brillouin zone points $\frac{m}{10}\boldsymbol{X}_i$ for $m = 0, 1, 2, 3, 4, 5$ and $i = 1, 2, 3$, where $\boldsymbol{X}_i$ represents the canonical unit vector for the i^{th} dimension. This selection of points is sufficient to approximate the bandgap because when the band structure of the $8\times8\times8$ supercell is "unwrapped" onto the Brillouin zone corresponding to a $1 \times 1 \times 1$ cell, the selected points unwrap onto the entire Brillouin zone. For a series of temperatures we calculate the full distribution of bandgaps, with the average and sample standard deviation of the calculated bandgaps at each temperature presented in Table 1.3.

T (K)	E_{gap} (eV)
220	0.422 ± 0.016
260	0.423 ± 0.019
300	0.432 ± 0.022
340	0.439 ± 0.024
380	0.446 ± 0.030
420	0.455 ± 0.038

Table 1.3: Temperature dependence of the average bandgap and spread of bandgaps. Bandgaps were calculated using instantaneous nuclear configurations obtained from MD simulations performed on an $8 \times 8 \times 8$ supercell at the various temperatures. The data are converged with respect to supercell size.

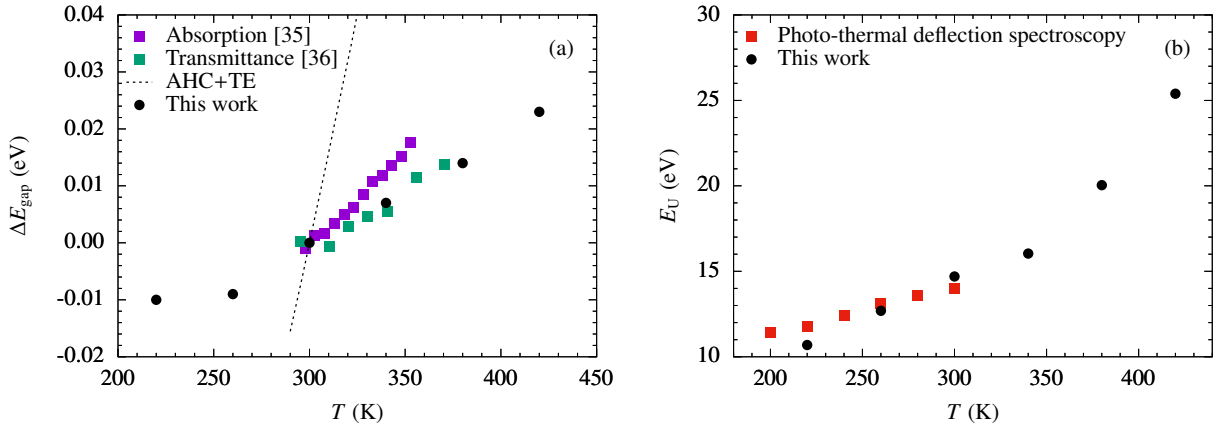

Figure 1.3: Characterization of (a) the temperature-dependent bandgap opening and (b) the temperature-dependent Urbach energy E_{U} compared in each case to experimental work. In (a), AHC+TE represents the standard Allen-Heine-Cardona estimate of vibrationally-induced bandgap change.

In Fig. 1.3a we compare the change of the bandgap relative to its value at 300K with experimental results for the same quantity. The model quite accurately captures the scale of bandgap opening in $MAPbI_3$. It should be noted that the contribution of temperature dependent changes to the mean lattice structure are less important in determining the bandgap opening than temperature dependent changes in the sampled regions of the potential energy landscape within our model. In Ref. [30] it was noted that higher-order electron-phonon coupling contributions are needed to accurately capture the temperature dependent bandgap opening in this system. In our theoretical approach these terms are naturally included via the non-linear nuclear coordinate dependence of the tight-binding parameters. Previous theoretical predictions (solid and dashed lines) both overestimated the magnitude of the bandgap opening.

A simple way to characterize the absorption edge is via knowledge of the width of the temperature-dependent bandgap distribution. Here we employ the simple model of Guerra *et al.* [31], which has successfully been used to fit the full absorption edge in $MAPbI_3$, including both the Urbach tail region and the onset of the edge. Within this model, the Urbach energy is related to the full width at half maximum of the bandgap distribution via

$$FWHM = 4E_{\mathrm{U}} \cosh^{-1}\sqrt{2}, \tag{1.3}$$

where within our model the full width at half maximum is estimated from the calculated temperature-dependent standard deviation of the bandgap distribution as reported in Table 1.3. Fig. 1.3b shows the calculated temperature dependence of the Urbach energy along with two data points from direct measurements taken from Ref. [32]. The agreement is remarkable.

1.4 Quantum Dynamics

To study charge-carrier dynamics, we employ a simplified version of mixed quantum-classical Ehrenfest dynamics. Within full Ehrenfest dynamics [33], the system Hamiltonian is written as a combination of electronic (quantum) and nuclear (classical) parts as

$$H = \hat{H}_{\text{TB}}(\boldsymbol{R}) + H_{\text{nuc}}(\boldsymbol{R}, \dot{\boldsymbol{R}}). \tag{1.4}$$

The atomic coordinates $\boldsymbol{R}$ are propagated classically using Hamilton's equations,

$$\begin{aligned} \frac{d\boldsymbol{R}}{dt} &= \dot{\boldsymbol{R}}, \\ \frac{d\dot{\boldsymbol{R}}}{dt} &= -\nabla_{\boldsymbol{R}}\Big(H_{\text{nuc}} + \langle \Psi(t)|\hat{H}_{\text{TB}}(\boldsymbol{R})|\Psi(t)\rangle \Big). \end{aligned} \tag{1.5}$$

For tight-binding studies of semiconductor mobilities like this one, the quantum wavefunction is normally propagated using first- or second-order Euler integration of the time-dependent Schrödinger equation:

$$i\hbar \frac{d\Psi(t)}{dt} = \hat{H}_{\text{TB}}(\boldsymbol{R}(t))\Psi(t). \tag{1.6}$$

Because the nuclear dynamics for the $MAPbI_3$ system are much slower than the electronic dynamics, our computational approach differs in two ways from this traditional approach. First, it is acceptable to make the approximation [34]

$$\frac{d\dot{\boldsymbol{R}}}{dt} \approx -\nabla_{\boldsymbol{R}} H_{\text{nuc}}, \tag{1.7}$$

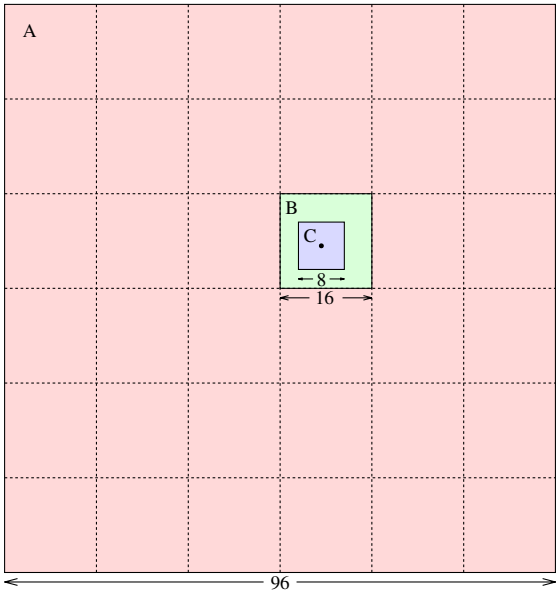

Figure 1.4: A summary of the periodic boundary conditions used in our model. Quantum dynamics are performed within the full supercell A, using matrix elements determined from a molecular dynamics trajectory performed in supercell B that are periodically replicated within supercell A. In calculating the on-site energy for an orbital of a particular atom (black dot in the diagram), an Ewald summation is performed using only atomic coordinates within supercell C centered on that atom.

to Eq. 1.5, meaning the nuclear coordinates feel only a very weak effect from the quantum charge carrier. This allows for the nuclear trajectories to be computed completely classically and subsequently fed into the quantum dynamical calculation. Second, we found it inefficient to use Euler integration of Eq 1.6. Instead, we utilize the more general truncated Taylor expansion to the quantum time-evolution operator:

$$\begin{aligned} \Psi(t+\Delta t) &= e^{-i\hat{H}_{\mathrm{TB}}(t)\Delta t/\hbar}\Psi(t) \\ &\approx \left(\sum_{n=0}^{C}\frac{(-i\Delta t/\hbar)^n}{n!}\hat{H}^n_{\mathrm{TB}}(t)\right)\Psi(t). \end{aligned} \tag{1.8}$$

For a timestep $\Delta t = 1$ fs, we found a cutoff of $C = 50$ to the sum in Eq. 1.8 sufficient to conserve the wavefunction norm to within a reasonable degree.

Nuclear configurations for a $16 \times 16 \times 16$ supercell were calculated at 1 fs intervals using molecular dynamics simulations within the canonical ensemble at $T = 220$ K, 260 K, 300 K, 340

K, 380 K, and 420 K [29]. The anisotropic lattice parameters for these simulations were calculated separately at each temperature using a calculation run in the isothermal-isobaric ensemble. The resulting nuclear configurations were used to calculate the time-dependent on-site and hopping matrix elements as described in Sec. 1.2.

On-site energies were calculated using only those atoms in the $8 \times 8 \times 8$ supercell centered on the relevant atom rather than the full $16 \times 16 \times 16$ MD supercell. It is possible that this procedure leads to unphysical atomic collisions on the edges of the $8 \times 8 \times 8$ subcell used for the Euler summation. However, we expect the motions of atoms $\gtrsim 4$ unit cells away from the relevant atom to largely cancel in the calculation of the on-site energy. Quantum dynamics were performed within a $96 \times 96 \times 96$ supercell consisting of six tiled $16 \times 16 \times 16$ MD supercells in each direction.

A summary of the different supercells used in our calculations is depicted in Fig. 1.4. We found that if supercell B were too small (i.e. $8 \times 8 \times 8$ or smaller instead of $16 \times 16 \times 16$), the calculated mobilities would be unphysically large due to finite size effects. The mobilities we report for a $16 \times 16 \times 16$ supercell B are the same as those that result from using a $25 \times 25 \times 25$ supercell B tiled within a $100 \times 100 \times 100$ supercell C.

1.5 Results: Electron and Hole Mobilities

All dynamics begin with a charge carrier localized to within about 1 nm. (A full discussion of the initial condition used is presented in the first Appendix to this chapter.) The charge spreads as time progresses, with the tails of its probability distribution eventually reaching the boundaries of the simulation supercell around $t = 150$ fs. Heat maps of the charge carrier density within a cross-section of the 3D simulation supercell are presented for $t = 5$ fs in Fig. 1.5a and for $t = 50$ fs in Fig. 1.5b. The charge spreads preferentially along the coordinate axes due to the spin-orbit coupling and the pp π couplings (which both allow right angle changes in the direction of the electron spreading) being smaller than the sp and pp σ hopping elements.

To compute the charge carrier mobility, we first calculate the thermally averaged spread of the

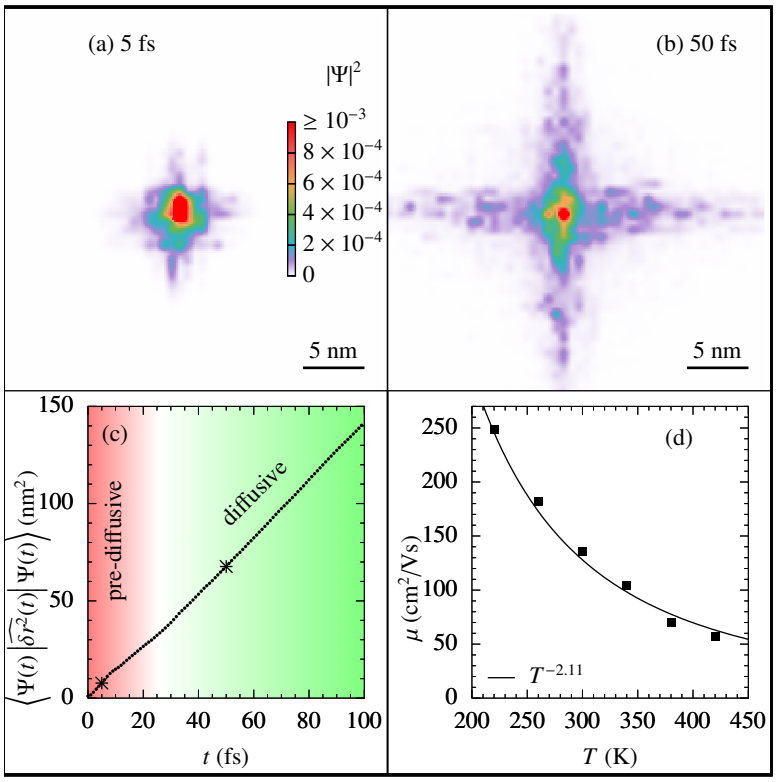

Figure 1.5: Carrier mobilities are calculated using the diffusion constant for spreading of the charge carrier. Heat maps for a single charge carrier density within a cross-section of the 3D system are shown at $t = 5$ fs in panel (a) and at $t = 50$ fs in panel (b). The simulation was initialized with a conduction band Wannier function, which is localized within about 1 nm. The spread of the carrier for this simulation is plotted vs. time in panel (c), with the data points corresponding to panels (a) and (b) highlighted. Using data averaged over multiple initial conditions, the temperature dependence of the mobility, plotted in panel (d), was found to most closely fit $T^{-2.11}$. Experiments suggest a temperature dependence of the form T^{-x} with $1.4 \lesssim x \lesssim 2.8$ [35, 36, 37, 38].

carrier as a function of time,

$$MSD(t) = \left\langle \frac{1}{Q} \mathrm{Tr}_{\mathrm{el}} \left\{ e^{-\beta \hat{H}_{\mathrm{TB}}(t=0)} \widehat{\delta r}^{\,2}(t) \right\} \right\rangle , \tag{1.9}$$

where $\langle \ldots \rangle$ denotes an average over classical trajectories, $\mathrm{Tr}_{\mathrm{el}}\{\ldots\}$ denotes a trace over electronic coordinates (either the valence band for hole mobilities or the conduction band for electron mobilities, obtained via projection onto the appropriate subspace), $Q = \mathrm{Tr}_{\mathrm{el}} e^{-\beta \hat{H}_{\mathrm{TB}}}$, and the action of $\widehat{\delta r}^{\,2}$

on a general wavefunction is described by

$$\begin{aligned} \widehat{\delta r}^{\,2}|\Psi(t)\rangle &= \left(\hat{r} - \langle\Psi(t)|\hat{r}|\Psi(t)\rangle\right)^2 \Big|\Psi(t)\Big\rangle, \\ \hat{r}|\boldsymbol{R}\rangle &= \boldsymbol{R}|\boldsymbol{R}\rangle. \end{aligned} \tag{1.10}$$

The mobility μ is then calculated from the diffusion coefficient for this spreading as

$$\begin{aligned} \mu &= \frac{e}{k_B T} \lim_{t\to\infty} \frac{MSD(t)}{6t} \\ &= \frac{e}{k_B T} D. \end{aligned} \tag{1.11}$$

In practice, it is numerically challenging to calculate $MSD(t)$ as defined in Eq. 1.9. It is simpler to instead use

$$MSD'(t) = \left\langle \frac{1}{N} \mathrm{Tr}_{\mathrm{el}} \left\{ \widehat{\delta r}^{\,2}(t) \right\} \right\rangle, \tag{1.12}$$

where N is the size of the electronic basis. Instead of performing a full trace, one can average over a representative subset of the full electronic basis via use of either valence Wannier functions (for hole mobilites) or conduction Wannier functions (for electron mobilities) centered on evenly spaced lattice sites. It has been argued for related models that such an approximation effectively captures the full thermal trace [34]. The effect on the calculated mobility of using Eq. 1.12 instead

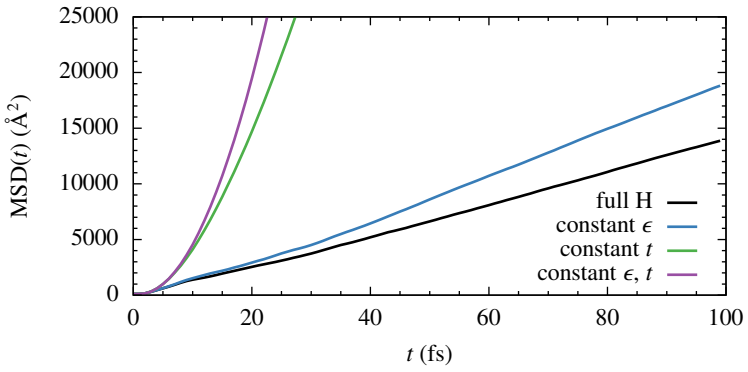

Figure 1.6: The function $\langle\Psi(t)|\widehat{\delta r}^{\,2}|\Psi(t)\rangle$, plotted for a single trajectory and initial electronic wavefunction. The wavefunction was propagated using different versions of $\hat{H}_{\mathrm{TB}}$.

of Eq. 1.9 is discussed in Sec. 1.7. A typical result for the quantity $\langle\Psi|\widehat{\delta r}^{\,2}(t)|\Psi\rangle$, the quantity which is averaged to obtain $MSD'(t)$, is plotted in Fig. 1.5c. The spreading reaches the diffusive (linear) regime fairly quickly; a reasonable estimate for the mobility can be inferred after just 10 fs.

In Fig. 1.6, we plot the same quantity $\langle\Psi|\widehat{\delta r}^{\,2}|\Psi\rangle$ for a single initial condition propagated using different versions of $\hat{H}_{\text{TB}}$. The black line uses the fully time-dependent tight binding Hamiltonian. The green (blue) line uses a version of that same Hamiltonian in which all off-diagonal (diagonal) matrix elements are set to their average value, but the diagonal (off-diagonal) matrix elements remain fully time-dependent. We measure the off-diagonal disorder to be the dominant factor in limiting the charge carrier mobility. The purple line uses a fully time-independent Hamiltonian, for which ballistic motion ($MSD(t) \propto t^2$) is observed at all times.

In Table 1.4, we list the average mobility corresponding to the spread of a conduction band wavefunction propagated using different versions of $\hat{H}_{\text{TB}}$ as described above. The effect of eliminating disorder in either the diagonal or off-diagonal part of the Hamiltonian is to increase the mobility, as expected. For this system, off-diagonal disorder is observed to play a greater role

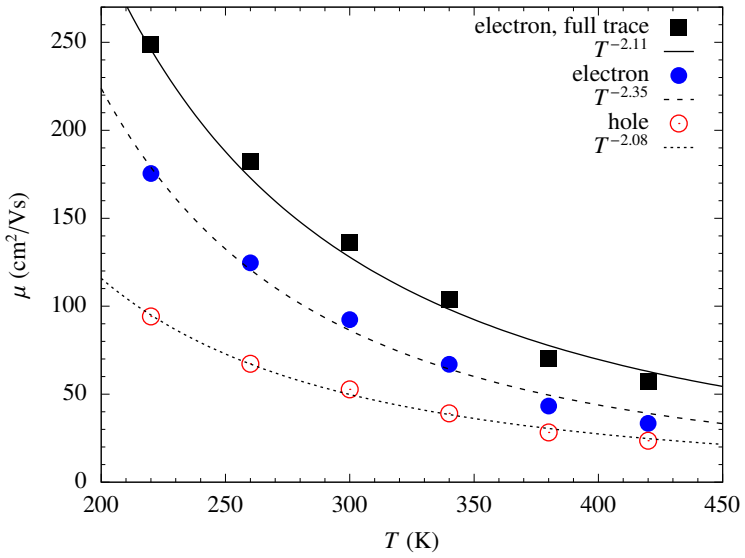

Figure 1.7: The charge carrier mobility, calculated using Eq. 1.9 for electrons and calculated using Eq. 1.12 for both electrons and holes. The data was fitted to a curve of the form aT^{-b}. The best estimate of the mobility of electrons at $T = 300$, calculated using the full thermal trace, is about 135 cm^2/Vs.

electronic Hamiltonian	μ (cm^2/Vs)
$\hat{H}_{\text{TB}}$	92
$\hat{H}_{\text{TB}}$ with all $\epsilon_i^\alpha(t) = \overline{\epsilon_i}$ constant	130
$\hat{H}_{\text{TB}}$ with all $t_{ij}^{\alpha\alpha}(t) = \overline{t_{ij}}$ constant	1100
$\hat{H}_{\text{TB}} - \hat{H}_{\text{SO}}$	32

Table 1.4: Calculated average mobilities for dynamics run with different versions of $\hat{H}_{\text{TB}}$.

than diagonal disorder in limiting the mobility, although including diagonal fluctuations generated simply from purely classical nuclear dynamics can underestimate polaron effects. When spin-orbit coupling is neglected from the calculation, the mobility is decreased by a factor of 3.

Mobilities calculated using the approximate version of the thermal charge spreading defined in Eq. 1.12 are plotted in Fig. 1.7 for electrons (blue closed circles) and holes (red open circles). The order of magnitude of the mobilities, approximately 100 cm^2/Vs, is similar to that observed in studies of large single crystals. Although some experimental studies have measured hole mobilities to be greater than electron mobilities by about five times, our results predict hole mobilities to be smaller than electron mobilities by about 30%, in line with some recent experimental work.

The temperature-dependent mobilities were fitted to a curve of the form $cT^{-\alpha}$. The study by Herz *et al*., which contains data spread across the largest temperature range, measures the electron mobility to vary as $T^{-1.5}$, which is the theoretical result for a system in which the dominant electron-phonon coupling is due to acoustic phonons. We predict the temperature dependence to be slightly greater, varying as $T^{-2.35}$ for electrons and $T^{-2.08}$ for holes. Recent experiments suggest a temperature dependence of $T^{-2.8}$ for electrons and $T^{-2.0}$ for holes [35]. If the calculation is done using the more exact thermal MSD function in Eq. 1.9, the temperature dependence is predicted to be $T^{-2.11}$ for electrons.

1.6 Conclusion

In this work, a tight-binding model was employed to study temperature-dependent absorption features and carrier mobilities and hybrid lead-halide perovskites. The model includes fluctuating diagonal (Holstein or Fröhlich) and off-diagonal (Peierls or Su-Schrieffer-Heeger) parts that were

parameterized using molecular dynamics and density functional calculations on small supercells. The temperature-dependent bandgap and the width of the absorption edge (Urbach energy) were calculated and found to agree quantitatively with available experimental data at all temperatures studied. Given the substantial disagreement within the experimental community over the magnitude and temperature dependence of the carrier mobilities, it is more difficult to assess the accuracy of the calculated carrier mobilities. However, their magnitude and temperature dependence fall within the accepted range and agree nearly quantitatively with one of the most recent experimental studies. The work points to the dominant role of the large off-diagonal (hopping) fluctuations in limiting the mobilities. Future work should focus on understanding the reason for long carrier lifetimes in these materials.

1.7 Appendix: Initial Conditions

1.7.1 Full thermal trace

We have also calculated the full thermal trace written in Eq. 1.9 in an approximate manner for negative charge carriers. This calculation is more demanding than the calculation of the more approximate $MSD'(t)$, and thus are statistics are worse. The electronic basis was comprised of the functions $|\uparrow;\boldsymbol{R}\rangle$, $|\downarrow;\boldsymbol{R}\rangle$, which represent the doubly-degenerate conduction band Wannier functions centered at each lattice vector $\boldsymbol{R}$ within the full $96 \times 96 \times 96$ quantum dynamics supercell. The action of the thermal operator could also be generated using a cutoff Taylor series as in Eq. 1.8. However, if the full thermal operator were applied directly to a basis state $|\phi\rangle$, low-energy eigenstates of $\hat{H}_{\text{TB}}(t=0)$, which exist in $|\phi\rangle$ in small magnitudes since $\hat{H}_{\text{TB}}(t=0)$ is in general disordered, would be amplified by about e^{400} with respect to the conduction band states that comprise $|\phi\rangle$. Therefore, it was necessary to apply the thermal operator to $|\phi\rangle$ in parts, each time projecting back onto the conduction band using the basis of conduction Wannier functions, i.e.

$$e^{-\beta\hat{H}_{\text{TB}}}|\phi\rangle = \left(\hat{P}e^{-\beta\hat{H}_{\text{TB}}/M}\right)^{M}|\phi\rangle, \tag{1.13}$$

where $\hat{P} = \sum_{\boldsymbol{R}} |\uparrow;\boldsymbol{R}\rangle\langle\uparrow;\boldsymbol{R}| + |\downarrow;\boldsymbol{R}\rangle\langle\downarrow;\boldsymbol{R}|$ and M was chosen to be 50.

A partial trace was performed at $T = 300$ K using just 8 basis functions and the resulting

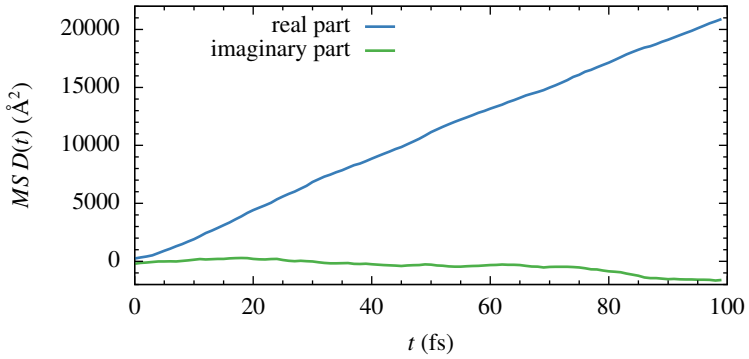

Figure 1.8: The real and imaginary parts of $MSD(t)$, calculated using a partial trace in Eq. 1.9.

$MSD(t)$ plotted in Fig. 1.8. The small observed imaginary part of $MSD(t)$ is expected to vanish in the full trace. The slope of the real part gives a mobility of 135 cm^2/Vs, compared to 92 cm^2/Vs when the Wannier initial conditions are used as in $MSD'(t)$.

1.7.2 Decomposition of full thermal trace

In this section, we present a mathematical motivation for using Eq. 1.12 instead of Eq. 1.9 in calculating the mobility and quantify the error in doing so. We do this to provide justification to the observation of Wang *et al.* who observed in a related, albeit simpler, model that the use of a localized initial condition may provide a good approximation to the full thermal trace [34]. The expression in Eq. 1.9 can be expanded by introducing a resolution of the identity operator:

$$\begin{aligned} MSD(t) &= \left\langle \frac{1}{Q} \sum_i \left\langle i \middle| e^{-\beta \hat{H}_{\text{TB}}(t=0)} \widehat{\delta r}^{\,2}(t) \middle| i \right\rangle \right\rangle \\ &= \left\langle \frac{1}{Q} \sum_{ij} \left\langle i \middle| e^{-\beta \hat{H}_{\text{TB}}(t=0)} \middle| j \right\rangle \left\langle j \middle| \widehat{\delta r}^{\,2}(t) \middle| i \right\rangle \right\rangle \\ &= \left\langle \frac{1}{Q} \sum_i \left\langle i \middle| e^{-\beta \hat{H}_{\text{TB}}(t=0)} \middle| i \right\rangle \left\langle i \middle| \widehat{\delta r}^{\,2}(t) \middle| i \right\rangle \right\rangle \\ &\quad + \left\langle \frac{1}{Q} \sum_{i \neq j} \left\langle i \middle| e^{-\beta \hat{H}_{\text{TB}}(t=0)} \middle| j \right\rangle \left\langle j \middle| \widehat{\delta r}^{\,2}(t) \middle| i \right\rangle \right\rangle, \end{aligned} \tag{1.14}$$

where all basis states are now indexed by i, j. For calculations in a model 1D system, the off-diagonal (latter) portion of this decomposition is observed to be small compared to diagonal (former) portion, and indeed our partial evaluation of the full thermal trace presented in Fig. 1.8 shows a similar result.

The diagonal part of Eq. 1.14 can be simplified in an approximate manner as follows:

$$\begin{aligned} \left\langle \frac{1}{Q} \sum_i \left\langle i \middle| e^{-\beta \hat{H}_{\text{TB}}(t=0)} \middle| i \right\rangle \left\langle i \middle| \widehat{\delta r}^{\,2}(t) \middle| i \right\rangle \right\rangle &\approx \left\langle \frac{1}{N} \sum_i \left\langle i \middle| \widehat{\delta r}^{\,2}(t) \middle| i \right\rangle \right\rangle \\ &\equiv MSD'(t), \end{aligned} \tag{1.15}$$

which is the $MSD'(t)$ defined in Eq. 1.12. Mobilities calculated in this manner are presented in Fig. 1.7. The mobilities were fitted to a curve of the form $T^{-\alpha}$. The exponent of the temperature dependence was found to be approximately 2.35 for electrons and 2.08 for holes. Approximate use of full thermal trace leads to similar results with somewhat larger mobility values and a slightly smaller exponent α.

1.8 Appendix: Matrix Element Spectral Densities and non-Gaussianity

1.8.1 Effective electron-phonon couplings and spectral coupling functions

If a first-order Taylor expansion of the tight-binding Hamitonian $\hat{H}_{\text{TB}}$ defined in Eq. 1.1 is made with respect to the deviations in the nuclear coordinates, the Hamiltonian becomes

$$\hat{H}_{\text{TB}} \approx \sum_{ij} \bar{t}_{ij}(\boldsymbol{R}) c_i^\dagger c_j + \sum_i \bar{\epsilon}_i(\boldsymbol{R}) c_i^\dagger c_i + \sum_{ijq} M_{ij}(q)(a_q + q_q^\dagger) c_i^\dagger c_j + \sum_{iq} N_i(q)(a_q + q_q^\dagger) c_i^\dagger c_i \quad (1.16)$$

As we have argued, such a small amplitude Taylor expansion likely fails for systems like $MAPbI_3$. However, if the relevant fluctuations $\delta t_{ij}^{\alpha\alpha}(\boldsymbol{R}(t))$ and $\delta\epsilon_i^\alpha(\boldsymbol{R}(t))$ are Gaussian, then such a Hamiltonian form holds with effective temperature-dependent electron-phonon coupling parameters [39].

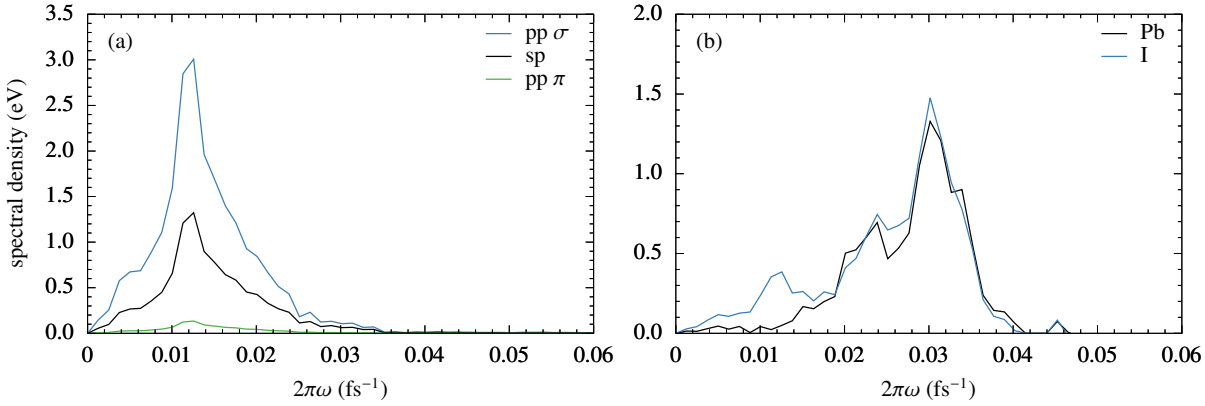

Figure 1.9: Spectral densities calculated at $T = 300$ K using the time-correlation function of the parameterized tight binding Hamiltonian matrix elements for (a) hopping amplitudes and (b) onsite matrix elements. In panel (b), the Pb s and Pb p on-site matrix elements yield identical spectral densities because the parameterized tight binding model assumes the fluctuations in those terms to be the same.

Contact is made with MD simulations by writing the electron-phonon couplings $M_{ij}(q)$ and $N_i(q)$ in terms of correlation functions. This is done via the introduction of spectral coupling functions. The spectral couplings Φ and Ξ are defined as

$$\begin{aligned} \Phi(\omega) &= \frac{2}{\hbar}\sum_q \delta(\omega-\omega_q)|M_{ij}(q)|^2 \\ \Xi(\omega) &= \frac{2}{\hbar}\sum_q \delta(\omega-\omega_q)|N_i(q)|^2. \end{aligned} \tag{1.17}$$

They are related to MD correlation functions by a cosine transform:

$$\langle \delta t_{ij}(t+s)\delta t_{ij}(s)\rangle_s = k_{\mathrm{B}}T\int d\omega \frac{\Phi(\omega)\cos(\omega t)}{\omega} \tag{1.18}$$

The same procedure can be performed for the spectral coupling of the onsite energy. Fig. 1.9 shows the computed spectral coupling for the hopping (s-p σ bond, p-p σ bond and p-p π bond) and the onsite energy (Pb s, Pb p and I p).

We emphasize again that the Hamiltonian form of Eq. 1.16 along with computed spectral functions that arise from the simulation of the cosine transform of the tight-binding parameter fluctuations does not imply standard linear coupling is valid, and can thus go beyond the harmonic approximation as long as fluctuations are Gaussian [39].

Thus the effective "electron-phonon" couplings will be temperature dependent and carry information related to the full non-linear behavior of the system. All that is needed for Eq. 1.16 to be valid is that the fluctuations are statistically Gaussian, which we assess in Sec. 1.8.2 below. It should be noted that the approach outlined here enables a fully quantum treatment of transport that goes beyond the Ehrenfest approach used in this work.

1.8.2 Non-Gaussian parameter

Here we assess if the tight binding parameters for our model of $MAPbI_3$ fluctuate in an essentially Gaussian manner. To do this we calculate the non-Gaussian parameter (NGP), which indicates deviations from Gaussian behavior. For a Gaussian distribution, this parameter will vanish. The NGP is defined as [40]

$$\text{NGP}(t) = \frac{\phi_4(t)}{3\phi_2^2(t)} - 1, \tag{1.19}$$

where $\phi_4 = \left\langle (h(s+t) - h(s))^4 \right\rangle$, $\phi_2 = \left\langle (h(s+t) - h(s))^2 \right\rangle$ and h can be either one of the three $t_{ij}^{\alpha\alpha}$ parameters or one of the ϵ_i^α parameters. Representative non-Gaussian parameters are plotted in Fig. 1.10. The NGPs associated with the on-site energies are clearly close to zero and certainly smaller than would be expected from, e.g., a telegraph process (data not shown). However, the NGPs for the hopping parameters appear to be larger, suggesting a model of the form 1.16 may not be quantitatively accurate for the treatment of $MAPbI_3$.

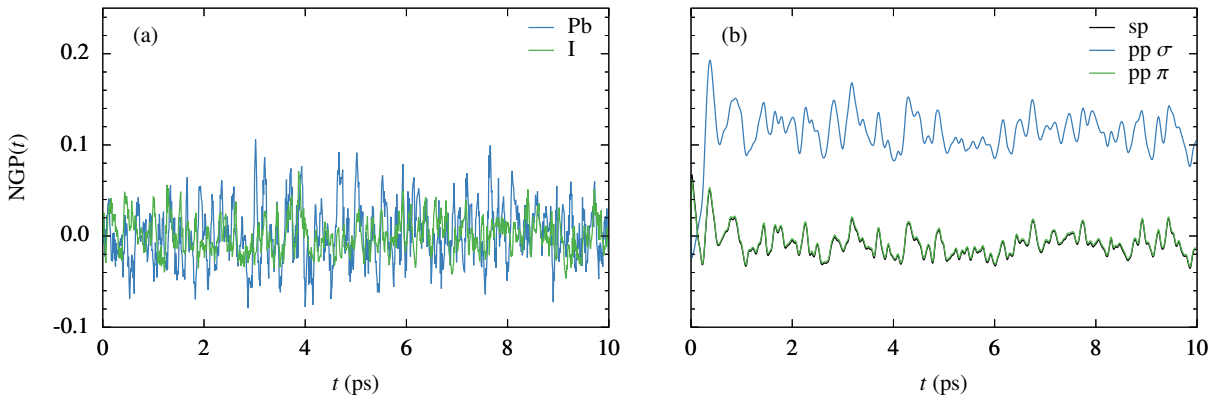

Figure 1.10: Calculated NGP for (a) on-site energies and (b) hopping amplitudes. In panel (a), the Pb s and Pb p on-site matrix elements yield identical non-Gaussian parameters because the parameterized tight binding model assumes the fluctuations in those terms to be the same.

CHAPTER 2
BINDING PROPERTIES OF SMALL CARRIER COMPLEXES IN TWO-DIMENSIONAL TRANSITION METAL DICHALCOGENIDES

2.1 Introduction

Atomically thin layers of crystalline transition metal dichalcogenides (TMDCs) have been the subject of intense investigation in recent years [41, 42]. As with graphene, TMDCs exhibit remarkable properties that originate from their quasi two-dimensional nature [43, 44]. Unlike graphene, however, single-layer TMDCs are direct gap semiconductors, opening up a wealth of potential practical applications ranging from field-effect transistors to photovoltaics [45, 46]. Furthermore, due to the lack of inversion symmetry in the quasi-two-dimensional version of these materials, the so-called K-points on opposite corners of the two-dimensional hexagonal Brillouin zone are inequivalent [47, 48, 49]. As a result, a distinct valley degree of freedom associated to states near these points emerges which may be manipulated and controlled, leading to the possibility of novel "valleytronic" applications [50, 51, 52, 53]. Lastly, carrier confinement to the material plane and reduced dielectric screening leads to large charge-carrier interaction effects, resulting in bound state complexes of electrons and holes with very large binding energies [54, 55, 56, 57, 58, 59, 60]. In this chapter we focus on this latter property, providing a deeper understanding of the factors that control the binding energies of electron-hole complexes in two-dimensional TMDCs.

From the computational perspective, the most accurate means of describing excitonic properties in periodic solids currently available is the GW+BSE approach [61, 62, 63, 64, 65]. Unfortunately, analogous fully *ab initio* approaches have not been developed for the treatment of larger electron-hole complexes such as trions and biexcitons [57, 58, 59, 60]. However, simplified approaches have proven to be effective, building on well-established coarse-grained methodologies developed for semiconductor quantum wells and other nanostructures. An effective real-space

electron-hole potential is combined with an approximate treatment of the band structure, such as an effective mass model or a few-band tight-binding model, to build the model Hamiltonian. This also has roots in the early discussion of the Bethe-Salpeter approach by Hanke and Sham, demonstrating the relationship to the phenomenological approach of Wannier [66].

The pioneering work of Keldysh highlighted the fact that screening effects in quasi two-dimensional systems are intrinsically non-local [67]. Using a generalized Keldysh approach, Cudazzo *et al.* formulated a simple and successful theory for excitons in graphane [68, 69]. This approach has since been applied by various authors to study optical spectra as well as the properties of excitons, trions, and biexcitons in TMDCs [70, 71, 72, 73, 74, 75]. These studies have produced exciton binding energies and real-space structures that are in reasonable quantitative agreement with first principles GW+BSE calculations and experiments in a variety of two-dimensional TMDC systems [76]. This fact is not entirely surprising for three reasons. First, recent *ab initio* calculations show that the effective quasiparticle interactions that emerge at the RPA level nearly perfectly match those used to describe the effective electron-hole interaction in the models mentioned above [77]. Second, the *ab initio* band structure near the K-point is well described by elementary two- and three-band models [52, 78]. Third, the spatial extent of the exciton that emerges from fully *ab initio* calculations is sufficiently large relative to the atomic scale to suggest that a coarse-grained Hamiltonian is justified [56].

Even within the simplified framework of an effective Hamiltonian, the exact solution of the multi-body Schrödinger equation for larger electron-hole complexes is challenging. Initial work on exciton and trion binding energies in TMDCs employed variational wave functions [70]. This approach has been used more recently and with more intricate trial wave functions to study biexcitons [57]. In both cases the results found from variational solutions of the effective few-body Schrödinger equation are in reasonable agreement with experimental results. However, since binding energies for trions and biexcitons are extracted with reference to the exciton binding energy, the use of variational wave functions for all excitonic complexes leads to binding energies that need not provide a lower bound to the "exact" value, and it is unclear how much error cancellation

occurs as a result. For the trion binding energy, Ganchev *et al.* have discovered a remarkable *exact* solution, but only for the case where the full Keldysh effective potential is replaced with a completely logarithmic form that is accurate only at short range [79].

The purpose of this work is to investigate the nature and accuracy of past estimations of charge carrier complex binding energies by comparing with numerically exact results, and thereby to provide insights into the properties of higher-order excitonic complexes in two-dimensional TMDCs. We begin by summarizing how electronic screening for two-dimensional materials differs from that of 3D ones, thus introducing the Keldysh potential. We then give an overview of the diffusion Monte Carlo method, which is used in this work to determine exact ground-state energies and radial distribution functions for two-, three-, and four-particle TMDC complexes. The results are compared to previous variational and experimental results, and the agreement is found to be quantitative. We conclude by discussing the potential for similar calculations to assist experimentalists in both understanding measurements of existing TMDC materials and in discovering similar materials with novel properties including high-temperature superfluidity.

2.2 Screening in two-dimensional transition metal dichalcogenides

The two-body Keldysh potential is

$$V(r) = \frac{\pi}{(\epsilon_1 + \epsilon_2) r_0} \left[H_0(r/r_0) - Y_0(r/r_0) \right], \tag{2.1}$$

where H_0 is the Struve function, Y_0 is the Bessel function of the second kind, and ϵ_1 and ϵ_2 are the dielectric constants for the material above and below the TMDC layer; in all results presented, we use $\epsilon_1 = \epsilon_2 = 1$, relevant for "ideal" or suspended TMDC monolayers.

Screening lengths and effective masses for all materials studied were determined in Ref. [70]. Potentials of the type 2.1 were first discussed by Keldysh [67] and have been used to treat excitonic properties in quasi two-dimensional materials in several recent studies [68, 69, 70, 71, 72, 73, 74, 75]. The potential 2.1 behaves as $1/r$ at large r/r_0, but diverges more weakly as $\ln(r/r_0)$ near

$r = 0$. The crossover point is related to the screening distance $r_0 = 2\pi\chi_{2D}$, where χ_{2D} is the two-dimensional polarizability of the TMDC layer. For computational efficiency and consistency with past variational calculations, we use a modified form of the effective potential 2.1, given by

$$V'(r) = -\frac{1}{r_0}\left[\ln\left(\frac{r}{r+r_0}\right) + (\gamma - \ln 2)e^{-r/r_0}\right], \tag{2.2}$$

where γ is Euler's constant [68]. Calculations with the unaltered potential 2.1 typically result in exciton ground-state energies that are merely 2–3 meV lower than those produced with 2.1.

2.3 Variational Monte Carlo

Variational Monte Carlo (VMC) is a conceptually simple method for estimating the energy corresponding to a trial wavefunction [80]. Whereas the average energy corresponding to a trial wavefunction can sometimes be computed analytically for low-dimensional systems, there is no hope for such an approach to succeed for higher-dimensional problems. Given a Hamiltonian $\hat{H}$ and a trial wavefunction $\Psi_T(\boldsymbol{R})$, the goal is to compute

$$\langle\Psi_T|\hat{H}|\Psi_T\rangle = \frac{\int d\boldsymbol{R}\Psi_T^*(\boldsymbol{R})\hat{H}\Psi_T(\boldsymbol{R})}{\int d\boldsymbol{R}|\Psi_T(\boldsymbol{R})|^2} \tag{2.3}$$

$$= \frac{\int d\boldsymbol{R}|\Psi_T(\boldsymbol{R})|^2 E_L(\boldsymbol{R})}{\int d\boldsymbol{R}|\Psi_T(\boldsymbol{R})|^2}, \tag{2.4}$$

where the *local energy* is defined as

$$E[\Psi_T] = E_L(\boldsymbol{R}) = \frac{\Psi_T^*(\boldsymbol{R})\hat{H}\Psi_T(\boldsymbol{R})}{|\Psi_T(\boldsymbol{R})|^2}. \tag{2.5}$$

The average energy of Eq. 2.4 can be seen as the weighted average of the local energy with respect to the weights $|\Psi_T(\boldsymbol{R})|^2$. The procedure for estimating this average within the Metropolis Monte Carlo scheme is straightforward. First, generate N points $\boldsymbol{R}_i$ that are distributed according to the probability distribution $|\Psi_T(\boldsymbol{R})|^2$. This is done by starting with a random $\boldsymbol{R}_1$. To generate $\boldsymbol{R}_{j+1}$ given $\boldsymbol{R}_j$, first generate a random candidate $\boldsymbol{R}'_{j+1}$. If $w = |\Psi_T(\boldsymbol{R}'_{j+1})|^2/|\Psi_T(\boldsymbol{R}_j)|^2 > 1$,

then set $\boldsymbol{R}_{j+1} = \boldsymbol{R}'_{j+1}$. Otherwise, a random number r is chosen from a uniform distribution on the interval $[0, 1]$. If $r < w$, then set $\boldsymbol{R}_{j+1} = \boldsymbol{R}'_{j+1}$. Otherwise, set $\boldsymbol{R}_{j+1} = \boldsymbol{R}_j$ again.

After the N points $\boldsymbol{R}_i$ have been chosen in this way, the estimate of the energy corresponding to the trial wavefunction is simply

$$E[\Psi_{\mathrm{T}}] \approx \frac{1}{N} \sum_{j=1}^{N} E_{\mathrm{L}}(\boldsymbol{R}_j). \tag{2.6}$$

In principle, the expectation value of any observable can be estimated in an analogous manner.

2.4 Diffusion Monte Carlo

In this section we outline the computational approach and model used to investigate excitonic complexes. Our technique of choice is diffusion Monte Carlo (DMC) [80]. The time-dependent Schödinger equation for a system of N electrons reads

$$i\hbar \frac{\partial \Psi(\boldsymbol{R}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \nabla^2_{\boldsymbol{R}} + V(\boldsymbol{R}) - E_{\text{ref}} \right] \Psi(\boldsymbol{R}, t), \tag{2.7}$$

where $\boldsymbol{R}$ is a $3N-$dimensional vector containing all electronic coordinates, $V(\boldsymbol{R})$ is the potential energy (usually considered to be pairwise additive), and we have made the assumption that any nuclei are fixed in space. A fixed references energy E_{ref} has been added for later use.

Let the eigenfunctions for the Hamiltonian in Eq. 2.7 be $\psi_j(\boldsymbol{R})$ and the corresponding eigenvalues be E_j for $j = 0, 1, 2, \ldots$ such that $E_j < E_k$ when $j < k$. Then any wavefunction $\Psi(\boldsymbol{R}, t)$ that satisfies Eq 2.7 can be written as a linear combination of these eigenfunctions that are each evolved in time according to their energy:

$$\Psi(\boldsymbol{R}, t) = \sum_j c_j \psi_j(\boldsymbol{R}) e^{-i(E_j - E_{\text{ref}})t/\hbar}, \tag{2.8}$$

where the coefficients c_j are defined by

$$c_j = \int d^{3N}\boldsymbol{R}\psi_j^*(\boldsymbol{R})\Psi(\boldsymbol{R}, t=0). \tag{2.9}$$

If $\Psi(\boldsymbol{R}, t)$ is propagated instead in *imaginary* time, its time development is altered slightly from Eq. 2.8 and now reads

$$\Psi(\boldsymbol{R}, t) = e^{-(E_0 - E_{\text{ref}})t/\hbar} \sum_j c_j \psi_j(\boldsymbol{R}) e^{-(E_j - E_0)t/\hbar}. \tag{2.10}$$

The effect of imaginary-time propagation is thus to exponentially project out all excited-state components of the initial wavefunction $\Psi(\boldsymbol{R}, 0)$, which then allows for computation of ground state properties, in particular the ground-state energy. Motivated by this result, DMC calculations aim to use an imaginary-time propagation of a sampling of the wavefunction Ψ in position space to obtain a sampling of the ground-state wavefunction ψ_0 and energy E_0.

The imaginary-time wavefunction can be written using the evolution operator as follows:

$$\begin{aligned}
\Psi(\boldsymbol{R}, it) &= e^{-t\hat{H}/\hbar}\Psi(\boldsymbol{R}, 0) && (2.11)\\
&= \left(e^{-t\hat{H}/N\hbar}\right)^N \Psi(\boldsymbol{R}, 0) && (2.12)\\
&= \left(e^{-t(\hat{V} - E_{\text{ref}})/N\hbar} e^{-t\hat{T}/N\hbar}\right)^N \Psi(\boldsymbol{R}, 0), && (2.13)
\end{aligned}$$

where in the final line a Trotter decomposition, which is only valid in the limit $N \to \infty$, has been used. In practice, rather than storing the entire wavefunction on a grid, the wavefunction is treated as a probability distribution and sampled by a set $\{\boldsymbol{R}_i^0\}_{1 \le i \le N_W}$ of N_W $3N$ dimensional vectors describing each electron coordinate called *walkers*. Of course, $\Psi(\boldsymbol{R}, t)$ is in general not always positive (or even real). To be able to treat $\Psi(\boldsymbol{R}, t)$ as a probability distribution, it must be assigned a nodal structure which is enforced in the sampling of walkers as a boundary condition: if a walker crosses the chosen nodal surface during the motions described below, it is eliminated to enforce zero walker density on the nodal surface. The method applied within this *fixed node*

approximation will yield the ground-state wavefunction with the assigned nodal structure. Its corresponding energy will be bounded from below by the actual ground state energy due to the variational principle.

The right-hand operator in Eq. 2.13 results in purely diffusive motion of the electronic coordinates because it represents the formal integration of the differential equation

$$-\hbar\frac{\partial}{\partial t}f(\boldsymbol{R},t) = -\frac{\hbar^2}{2m}\nabla^2 f(\boldsymbol{R},t) \tag{2.14}$$

by a time t/N. Thus, if the function f is sampled at time t_0 by a set $\{\boldsymbol{R}_i^0\}_{1\le i\le N_W}$ of N_W walkers, then the sampling of f at time $t_0 + t/N$ becomes $\{\boldsymbol{R}_i^0 + \sqrt{\hbar t/2mN}\Delta g\}$, where Δg is a $3N$-dimensional vector of gaussian random numbers with variance 1.

The left-hand operator in Eq. 2.13 is the formal integration of the first-order differential equation

$$-\hbar\frac{\partial}{\partial t}f(\boldsymbol{R},t) = (V(\boldsymbol{R}) - E_{\text{ref}})f(\boldsymbol{R},t). \tag{2.15}$$

This equation describes exponential decay, as in a first-order chemical reaction. Walker density at the point $\boldsymbol{R}$ either increases or decreases at a rate corresponding to the value $V(\boldsymbol{R}) - E_{\text{ref}}$. Thus, given a sampling of f by N_W walkers $\{\boldsymbol{R}_i^0\}_{1\le i\le N_W}$ at time t_0, a sampling of f after propagation by the operator $e^{-t\hat{V}/N\hbar}$ can be obtained by either eliminating or duplicating the walker $\boldsymbol{R}_i$ with a probability $\left|(V(\boldsymbol{R}) - E_{\text{ref}})\frac{t}{N}\right|$. The walker is eliminated with this probability if $V(\boldsymbol{R}) - E_{\text{ref}} > 0$ and is duplicated with that probability otherwise. (In another incarnation of the method, each of the N_W walkers carries a weight W_i and instead of walkers being eliminated or duplicated in this step, each walker's weight is multiplied by the factor $\exp(-t\hat{V}(\boldsymbol{R}_i)/N\hbar)$. This version of the method is usually more expensive because walkers with very small weights, i.e. little effect on the converged energy, require the same amount of computational effort to propagate.)

The ground state energy is estimated at each step in the propagation as

$$E_0 \approx \frac{1}{N_W}\sum_i V(\boldsymbol{R}_i). \tag{2.16}$$

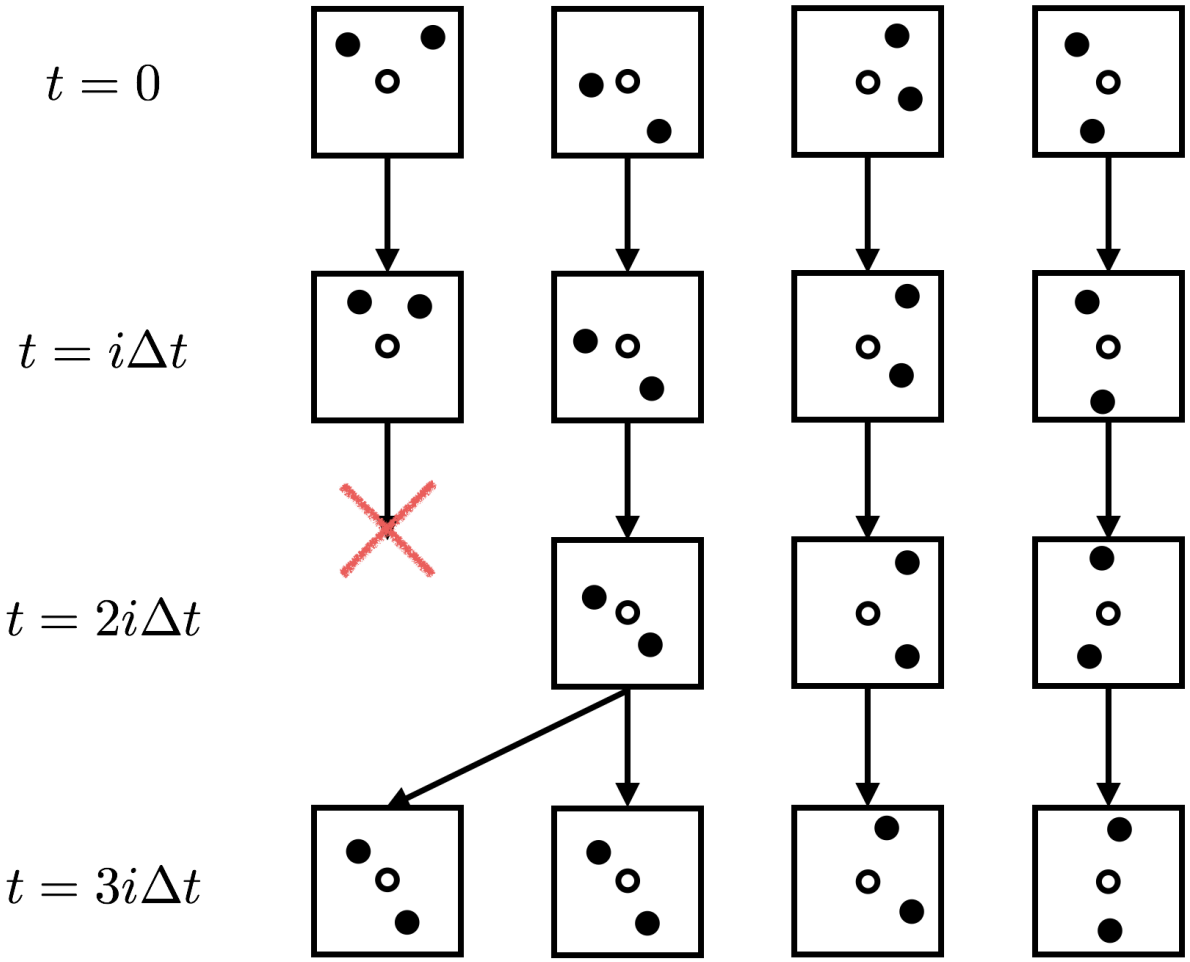

Figure 2.1: A schematic depiction of the DMC imaginary time propagation process for the calculation of the ground state trion energy. Each square corresponds to a single walker at a point in imaginary time. Electrons (holes) are depicted with filled (unfilled) circles. As only the relative electron and hole coordinates are significant, the hole is always depicted in the same position.

In practice, the reference energy E_{ref} is altered after each propagation step to enforce a stable walker population. More specifically, this work uses the reference energy update

$$E_{\text{ref}} = \frac{1}{N_W} \sum_i V(\boldsymbol{R}_i) + c(N_W^0 - N_W), \tag{2.17}$$

where N_W^0 is the desired number of walkers (chosen with consideration of the computational resources available) and c is chosen from test runs to be a reasonable constant. This definition leads to a slightly higher chance of particle duplication during the next propagation step if $N_W < N_W^0$

and vice-versa, thereby stabilizing the walker population.

The imaginary time propagation process is depicted schematically in Fig. 2.1 for a calculation of the ground state trion energy. Each square corresponds to a single walker at a point in imaginary time. Electrons (holes) are depicted using filled (unfilled) circles. In the first timestep, all four walkers diffuse and none are eliminated or duplicated. In the next timestep, a walker with high potential energy compared to the others (due to the electrons being in close vicinity to one another) is destroyed. In the following timestep, a walker with low potential energy compared to the others is duplicated. This process is continued indefinitely until the average ground state energy reaches a specified degree of accuracy. Because walker positions at successive imaginary times are highly correlated, a blocking procedure must be used when averaging so as to not underestimate the error in the average energy.

As outlined above, DMC will yield the correct ground-state energy (within the fixed-node approximation). However, the energy will converge poorly due to the $1/r$ divergence of the interparticle interaction. To remedy this drawback, a mixed probability distribution $f(\boldsymbol{R},t) = \Psi_{\mathrm{T}}(\boldsymbol{R})\Phi(\boldsymbol{R},t)$ is more commonly used [81, 82]. The function Ψ_{T}, sometimes called a *guiding wavefunction*, enhances sampling by providing exactly the correct divergence in its kinetic energy to balance the divergence in the potential energy. The imaginary time Schrödinger equation for this new probability distribution becomes

$$-\frac{\partial f(\boldsymbol{R},t)}{\partial t} = -\frac{1}{2}\nabla^2_{\boldsymbol{R}} f(\boldsymbol{R},t) + \nabla_{\boldsymbol{R}}\cdot[\boldsymbol{v}(\boldsymbol{R})f(\boldsymbol{R},t)] + (E_{\mathrm{L}}(\boldsymbol{R}) - E_{\mathrm{ref}})f(\boldsymbol{R},t), \tag{2.18}$$

where

$$\boldsymbol{v}(\boldsymbol{R},t) = \Psi_{\mathrm{T}}^{-1}(\boldsymbol{R})\nabla_{\boldsymbol{R}}\Psi_{\mathrm{T}}(\boldsymbol{R}), \tag{2.19}$$

$$E_{\mathrm{L}}(\boldsymbol{R}) = \Psi_{\mathrm{T}}(\boldsymbol{R})^{-1}\hat{H}\Psi_{\mathrm{T}}(\boldsymbol{R}), \tag{2.20}$$

where $\boldsymbol{v}$ is called the electronic drift and E_{L} is called the local energy. The algorithm for propagating the walkers remains the same as that outlined above (replacing $V(\boldsymbol{R})$ with $E_{\mathrm{L}}(\boldsymbol{R})$), except that

additional drift process is carried out. After a walker diffuses and before it is possibly duplicated or eliminated, the walker's position moves from $\boldsymbol{R}_i$ to $\boldsymbol{R}_i + \boldsymbol{v}(\boldsymbol{R}_i)$. This importance-sampled DMC variant results in walkers which converge to the mixed probability distribution $\Psi_{\mathrm{T}}\Psi$, where Ψ is the exact solution to Schrödinger's equation *sharing the nodal structure of* Ψ_{T}.

Given that the only variable information needed for the propagation of each individual walker is the reference energy E_{ref}, the DMC method is straightforward to parallelize, similar to other quantum Monte Carlo approaches. In fact, the optimization of graphical processing units (GPUs) for linear algebra manipulations makes them ideal candidates for performing DMC calculations for systems of dozens of electrons because of the matrix determinants required to enforce the fixed-node approximation.

For the systems we consider, the exact ground state wave function is nodeless because the complexed studied contain a maximum of two electrons and two holes. Therefore DMC yields *exact* ground state energies and a sampling of the *exact* ground state wavefunctions. The guiding wavefunction we use is of the form $\Psi_{\mathrm{T}}(\boldsymbol{R}) = e^{J(\boldsymbol{R})}$, which contains the Jastrow factor introduced in Ref. [79] adapted specifically to the potential 2.1. The Jastrow term contains two-body electron-hole and electron-electron (hole-hole) terms

$$u_{\mathrm{eh}}(r) = c_1 r^2 \ln(r) e^{-c_2 r^2} - c_3 r \left(1 - e^{-c_2 r^2}\right), \tag{2.21}$$

$$u_{\mathrm{ee}}(r) = c_4 r^2 \ln(r) e^{-c_5 r^2}. \tag{2.22}$$

The constants $c_1 = m_{\mathrm{e}} m_{\mathrm{h}}/2(m_{\mathrm{e}} + m_{\mathrm{h}})$ and $c_4 = -m_{\mathrm{e}}/4$, where $m_{\mathrm{e,h}}$ are the effective masses of the carriers, were chosen to satisfy the logarithmic analogue of the Kato cusp conditions.

The other parameters in the Jastrow terms can be chosen to maximize sampling efficiency. Naïvely, one might propose choosing the parameters so as to minimize the energy as computed using VMC. However, they can be more efficiently optimized using the process of *unreweighted variance minimization* [83, 84]. For a given guiding wavefunction $\Psi_{\mathrm{T}}^{\{\alpha_0\}}$ which is a function of the

set of chosen parameters $\{\alpha_0\}$ and a set of N points $\boldsymbol{R}_i$ distributed according to the probability distribution $|\Psi_{\mathrm{T}}|^2$, the variance in the energy can be estimated as

$$\sigma_{\mathrm{E}}^2 = \frac{1}{N}\sum_{j=1}^{N}\left(E_{\mathrm{L}}^{\{\alpha_0\}}(\boldsymbol{R}_i) - E[\Psi_{\mathrm{T}}^{\{\alpha_0\}}]\right)^2, \tag{2.23}$$

where the estimate in the energy $E[\Psi_{\mathrm{T}}^{\{\alpha_0\}}]$ is the same as in Eq. 2.6. For the exact ground state wavefunction Ψ_0, the local energy will be constant and the variance in the energy estimate will be zero. Thus, the variance in the energy estimate can be used as a measure of the wavefunction quality; in general, parameter sets which result in a lower variance in the energy estimate correspond to better guiding wavefunctions.

The guiding wavefunction parameters are optimized via a correlated sampling approach which uses the points $\boldsymbol{R}_i$ generated according to the probability distribution $|\Psi_{\mathrm{T}}^{\{\alpha_0\}}|^2$. For a new set of parameters $\{\alpha\}$, the *reweighted* variance (i.e. estimate of the true variance) in the energy estimate will be

$$\sigma_w^2 = \frac{1}{\sum_{j=1}^{N} w_{\{\alpha_0\}}^{\{\alpha\}}(\boldsymbol{R}_j)}\sum_{j=1}^{N} w_{\{\alpha_0\}}^{\{\alpha\}}(\boldsymbol{R}_j)\left(E_{\mathrm{L}}^{\{\alpha\}}(\boldsymbol{R}_j) - E_{\mathrm{r}}[\Psi_{\mathrm{T}}^{\{\alpha\}}]\right)^2, \tag{2.24}$$

where the reweighted energy is

$$E_{\mathrm{r}}[\Psi_{\mathrm{T}}^{\{\alpha\}}] = \frac{1}{\sum_{j=1}^{N} w_{\{\alpha_0\}}^{\{\alpha\}}(\boldsymbol{R}_j)}\sum_{j=1}^{N} w_{\{\alpha_0\}}^{\{\alpha\}}(\boldsymbol{R}_i)E_{\mathrm{L}}^{\{\alpha\}}(\boldsymbol{R}_j) \tag{2.25}$$

and the factor

$$w_{\{\alpha_0\}}^{\{\alpha\}}(\boldsymbol{R}_i) = \frac{|\Psi_T^{\{\alpha\}}(\boldsymbol{R}_i)|^2}{|\Psi_T^{\{\alpha_0\}}(\boldsymbol{R}_i)|^2} \tag{2.26}$$

is used to reweight the probability distribution so that it corresponds to the new guiding wavefunction $\Psi_{\mathrm{T}}^{\{\alpha\}}$.

	DMC X (eV)	variational X	DMC X^- (meV)	experimental X^-	variational X^-
MoS_2	0.5514	0.54	33.8	43 [86], 18 [58]	26
$MoSe_2$	0.4778	0.47	28.4	30 [87]	21
WS_2	0.5191	0.50	34.0	30 [88], 45 [89]	26
WSe_2	0.4667	0.45	29.5	30 [90]	22

Table 2.1: Estimated exciton (X) and trion (X^-) binding energies for different members of the 2D TMDC class of materials. Where two numbers are reported, the number on the left is the most current estimate. The statistical uncertainty in the DMC data is on the order of 0.1–0.3 meV. The column labeled "variational" refers to results based on the Keldysh form, taken from Ref. [70].

	DMC XX (meV)	experimental XX
MoS_2	22.9	70 [91]
$MoSe_2$	18.4	
WS_2	24.1	65 [88]
WSe_2	20.9	52 [57]

Table 2.2: Estimated biexciton (XX) binding energies for different members of the 2D TMDC class of materials. The statistical uncertainty in the DMC data is on the order of 0.1–0.3 meV.

The *unreweighted* variance is just

$$\sigma_u^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left(E_{\mathrm{L}}^{\{\alpha\}}(\boldsymbol{R}_j) - E[\Psi_{\mathrm{T}}^{\{\alpha\}}]\right)^2. \tag{2.27}$$

This is usually a more stable cost function to use for guiding wavefunction parameter optimization, especially for large systems where small changes in the parameters can lead to large reweighting factors $w_{\{\alpha_0\}}^{\{\alpha\}}(\boldsymbol{R}_i)$.

Finally, a blocking method is used to gauge the correlation timescales for the energy estimates and yields accurate standard deviations for the final average [85].

2.5 Results

Energy estimates were obtained for calculations with $\Delta t \in \{0.01, 0.03, 0.1\}$ and then extrapolated to zero timestep. DMC probability distributions were sampled from forward-walking DMC calculations [92] with the optimal guiding wave function described by Eqs. (2.21)–(2.22), and $\Delta t = 0.01$.

At convergence, DMC yields numerically exact exciton, trion and biexciton ground-state energies within the confines of an effective few-body Schrödinger equation. Specifically, our calculations employ an effective mass treatment of the band structure and a screened Coulomb interaction appropriate for the two-dimensional TMDC family of materials, i.e.

$$H = -\sum_i \frac{\nabla_i^2}{2m_i} + \sum_{i<j} q_i q_j V(r_{ij}). \tag{2.28}$$

Hamiltonians of this form have been successfully used to describe excitons, trions, and biexcitons in semiconductor quantum wells [93, 94, 95], as well as scenarios such as excitons in doped quantum wells and electron-hole plasmas [96]. Note that Eq. (1) neglects possible three-body (and higher-order) effective screened Coulomb interactions between carriers; we will evaluate the success of this approach for monolayer TMDCs and return to this point before concluding.

In addition, DMC allows a full sampling of the square of the wavefunction, which can be used to extract insight into the structure of small bound carrier assemblies. Although DMC has previously been used to calculate ground-state properties for trions interacting with a purely logarithmic potential [79], to the best of our knowledge it has not been used to calculate trion properties with the more realistic electron-hole interaction above, nor has it been used to calculate the properties of biexcitons. It should be noted that while the present work was underway, a numerically exact finite temperature path integral Monte Carlo (PIMC) study of excitons, trions, and biexcitons using the full Keldysh effective potential appeared [97]. While we believe that DMC is a more direct method than PIMC for the study of what are essentially ground state properties, we note that the results presented here are in quantitative agreement with those presented earlier in Ref. [97], yielding ground state energies that lie below those of Ref. [97] by fractions of a percent. On the other hand, the goals of this work are somewhat distinct from those of Ref. [97]. In particular, we focus on the specific physical factors that influence the delicate balance of relative trion and biexciton binding energies, as well as the accuracy of variational approaches in light of the "exact" DMC results.

In Tab. 2.1, we report exciton and trion binding energies calculated via DMC and compare to

	Keldysh	pure ln	pure $1/r$	variational 1	variational 2
MoS_2	33.8	48.9	1630	26	14
$MoSe_2$	28.2	39.3	1780	21	12
WS_2	33.9	53.6	1050	26	14
WSe_2	29.5	44.9	1120	22	12

Table 2.3: Comparison of trion binding energies for several potential forms in units of meV obtained from DMC, except for the column labeled "variational"; the latter results are based on the Keldysh form and taken from Ref. [70]. Binding energies in the "variational 1" column are with respect to the variational exciton binding energies, whereas those in the "variational 2" column are with respect to the exact DMC exciton binding energies. For the DMC Keldysh and pure logarithmic potentials, the uncertainty is of the order of 0.1–0.3 meV. For the pure Coulombic potential, the uncertainty is of the order of 10 meV.

those extracted in recent experiments. The DMC exciton binding energies, defined as $E_{\mathrm{b}}^{\mathrm{X}} = -E^{\mathrm{X}}$, are only 2–4% larger than those obtained in previous variational calculations that employed a trial wavefunction of the form $\Psi_{\mathrm{T,var}}(r_{\mathrm{eh}}) \sim \exp(-r_{\mathrm{eh}}/a)$. The radial probability distribution for the distance r_{eh}, which completely determines the exciton wavefunction, is plotted in Fig. 2.2(a); we compare the variational wavefunction to results obtained via DMC as well as a grid-based exact diagonalization of the one-dimensional Schödinger equation. The simple $1s$-like variational wave-

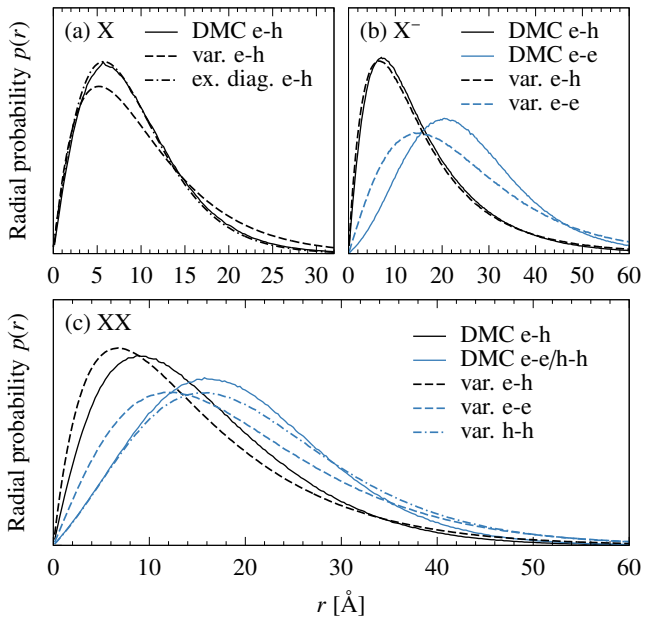

Figure 2.2: (a) Radial probability distributions for the distance r_{eh} of an exciton in MoS_2. (b) Radial probability distributions for the distances r_{eh} and r_{ee} of a negative trion in MoS_2. (c) Radial probability distributions for the distances r_{eh}, r_{ee}, and r_{hh} of a biexciton in WSe_2. For the DMC calculation, the curves for r_{ee} and r_{hh} coincide because the electron and hole effective masses are taken to be equal.

	Keldysh	pure ln	pure $1/r$	variational 1	variational 2
MoS_2	21.3	26.2	2610		
$MoSe_2$	17.0	21.2	2820		
WS_2	22.6	28.7	1680		
WSe_2	19.5	23.9	1780	37	16

Table 2.4: Comparison of biexciton binding energies for several potential forms in units of meV obtained from DMC, except for the column labeled "variational"; the latter results are based on the Keldysh form and taken from Ref. [57] for the WSe_2 biexciton. Binding energies in the "variational 1" column are with respect to the variational exciton binding energies, whereas those in the "variational 2" column are with respect to the exact DMC exciton binding energies. For the DMC Keldysh and pure logarithmic potentials, the uncertainty is of the order of 0.1–0.3 meV. For the pure Coulombic potential, the uncertainty is of the order of 10 meV.

function matches the true ground-state wavefunction well, but does not decay rapidly enough for large r. As we will show, achieving a similar level of agreement between exact DMC and variational estimates for the binding energies of larger excitonic complexes is, in principle, a much more difficult task because trion and biexciton wave functions are more elaborate and their approximation may in principle require many variational parameters to achieve a high level of accuracy.

The DMC trion binding energies given in Tab. 2.1 are all in the range of 28–34 meV, which is in excellent agreement with current experimental estimates; however it should be noted that realistic substrate effects have been ignored in the present calculations. In Tab. 2.3, we compare trion binding energies for two additional potentials: a purely logarithmic form and an unscreened $1/r$ Coulombic form. These two potentials represent the asymptotic small and large r behavior, respectively, of Eq. 2.1. The purely logarithmic potential approximation has been employed by Ganchev *et al.* in their analytical treatment of trions in TMDCs [79] while the Coulomb potential is the standard form for three-dimensional semiconductors. We find that a purely logarithmic potential overbinds the trion and results in binding energies about 50% larger than those reported by experiments. Unsurprisingly, the pure Coulombic potential vastly overbinds the complex, resulting in binding energies that are 30–50 times too large and with a different ordering than is the case for the full potential Eq. 2.1, which is material dependent. Coulombic binding energies would of course be reduced with the inclusion of a static dielectric constant significantly larger than unity.

Our trion binding energies are about 30% larger than those calculated variationally. The two-

parameter variational trial wavefunction used in the trion calculations was [70]

$$\Psi_{\mathrm{T,var}}(r_{\mathrm{e_1h}}, r_{\mathrm{e_2h}}) \sim \exp\left(-r_{\mathrm{e_1h}}/a - r_{\mathrm{e_2h}}/b\right) + \{a \leftrightarrow b\}, \tag{2.29}$$

a form inspired by Chandrasekhar's treatment of the hydrogen anion [98]. Although Fig. 2.2(b) shows that this optimized variational wavefunction reproduces $p(r_{\mathrm{eh}})$ almost exactly, the variational form does not capture the electron-electron repulsion properly because it lacks any explicit $r_{\mathrm{e_1}} - r_{\mathrm{e_2}}$ correlation terms. The peak of the electron-electron distribution is at too small a radius, which results in over-estimating the electron-electron repulsion and underestimating the trion binding energy, as seen in Tab. 2.3. Nonetheless, the level of agreement is surprisingly good given the simplicity of the variational wave function employed in Ref. [70]. Furthermore, by treating the exciton and trion on equal footing with physically similar variational wavefunctions, a fortuitous cancellation of total energy errors leads to binding energies which are quite close to the exact results ("variational 1" column in Tab. 2.3); referencing the variational trion energy to the exact DMC exciton energy ("variational 2" column in Tab. 2.3) leads to a significant underestimation of the binding energy, albeit one that is a genuine lower bound.

Finally, in Tabs. 2.2 and 2.4 we report biexciton binding energies $E_{\mathrm{b}}^{\mathrm{XX}} = 2E^{\mathrm{X}} - E^{\mathrm{XX}}$, and in Fig. 2.2(c) we compare carrier probability distributions obtained from DMC and from a recent six-parameter variational calculation [57]. Other than the r_{hh} distance, which is accurately predicted, the variational wavefunction is a bit too compact, leading to a total variational energy which is slightly too high. When referenced to the exact DMC exciton energy, we note that the variational biexciton binding energy is actually quite accurate – only 3.5 meV (about 15%) too small. When referenced to the *less accurate* variational exciton energy, the mis-matched cancellation of errors is such that the biexciton binding energy is slightly *overestimated*.

In our DMC calculations, we find that the full potential with dielectric screening Eq. 2.1 yields binding energies that are smaller than either of the other potentials considered, and that are smaller than experimental estimates by 60–70%. Most significantly, exact DMC calculations show that the

binding energies for biexcitons are significantly *smaller* than that of trions. When this work was initially published, this fact, which was noted in PIMC calculations as well [97], disagreed with the best existing experimental estimates and was at odds with expectations that emerge from the standard case of pure Coulombic interactions: whereas in the latter $1/r$ case biexcitons are more strongly bound than trions by a factor of about 1.6, in the purely logarithmic case the situation is reversed and trions are more strongly bound than biexcitons. Interestingly, we find that for realistically parameterized Keldysh potentials, the biexciton binding energies are slightly smaller than those found with the purely logarithmic potential, despite the latter being a presumably "weaker" potential; this highlights the subtle balance of energies involved in the formation of the biexciton. We note in passing that the binding energies for biexcitons obtained with the logarithmic potential are significantly closer to the full Keldysh results than they are for trions, suggesting that the short-range approximation of Ganchev *et al.* may be even better for biexcitons. This result is consistent with the smaller real-space structure of the biexciton seen by comparing Figs. 2.2(b) and (c).

To gain deeper insight into this balance of energies, we consider the electron-hole and electron-

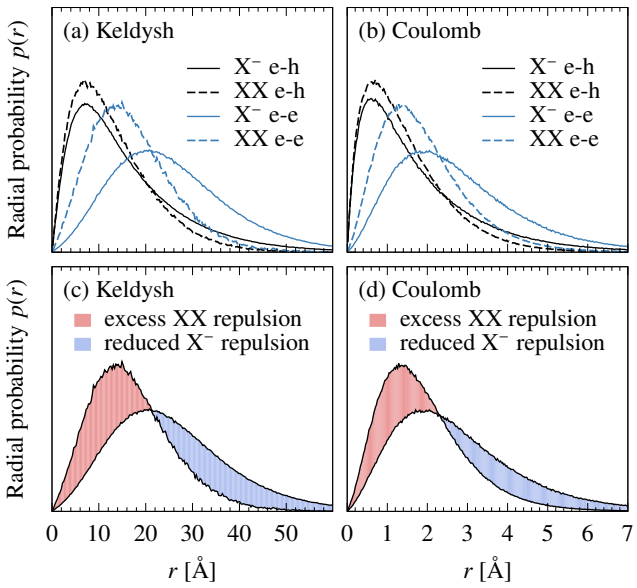

Figure 2.3: (a) Radial probability distributions for the distances r_{eh} and r_{ee} in the trion (X^-) and biexciton (XX) using a Keldysh form for the inter-carrier potential. (c) The same r_{ee} distributions as in panel (a), but the relocated repulsive weight is shaded. (b),(d) The same as in panels (a),(c) but using a Coulombic form for the inter-carrier potential.

electron distributions for trions and biexcitons obtained with the Keldysh and Coulombic potentials, plotted in Fig. 2.3(a),(b). Suppose that for a given potential, the attractive electron-hole probability profiles were identical for the trion and the biexciton, and the repulsive electron-electron (hole-hole) profiles were also identical for the trion and the biexciton. Then elementary arguments using the definition of the trion and biexciton binding energies, along with the pairwise additive potential, show that the biexciton binding energy would be exactly twice the trion binding energy, $E_{\mathrm{b}}^{\mathrm{XX}}/E_{\mathrm{b}}^{\mathrm{X}^-} = 2$. Any deviations from this ideal ratio are due to *relative* differences in the attractive and repulsive probability distributions as the second hole is added to the negative trion.

Instead, biexciton-to-trion binding energy ratios of less than 2 are observed for both the screened interaction Eq. 2.1 and the Coulomb interaction. In both cases, about 9% of the total weight in the $p(r_{\mathrm{eh}})$ (attractive) profile is relocated from long r to short r. More significantly, a much larger fraction of the total weight in the $p(r_{\mathrm{ee}})$ (repulsive) is relocated to short r, leading to a reduced biexciton-to-trion ratio. Specifically, for the screened potential Eq. 2.1, about 31% of the weight is relocated, which leads to a biexciton binding energy that is smaller than the trion binding energy; for the Coulomb potential, only 26% of the weight is relocated, and the biexciton binding energy remains larger than the trion binding energy. The relocated repulsive area is shaded for both potentials in Fig. 2.3(c),(d). From an energetic standpoint, this more notable change in the repulsive profile occurs because in the trion, there is no reason for the like charges to be physically close in space. However, in the biexciton, the complex can achieve stabilizing electron-hole interactions by having the like charges closer together in space.

As noted above, our calculated biexciton binding energies differ substantially from most experimental estimates of the biexciton binding energy. There are a number of possible sources of error in our calculation, including the neglect of substrate screening and higher-order charge-carrier interactions. On the other hand, experimental determination of biexciton binding energies in the TMDCs is quite difficult and involves both assumptions of the nature of spectral signals as well as extrapolations. Indeed, after this work was published and while this thesis was written, a preprint appeared which provides an update to this interesting discrepancy [99]. The authors attribute the

difficulty in assigning the biexciton peak to inhomogeneous broadening and the presence of defect bands. They exploit recent advances in device fabrication to suppress these unwanted effects and were able to unambiguously reassign the previously assigned biexciton peak to the charged five-member complex XX^-. The new biexciton peak yields a WSe_2 biexciton binding energy of approximately 18 meV, in agreement with our calculated result.

2.6 Conclusion

In this work, small ground-state carrier complexes in 2D transition metal dichalcogenides were studied within a real-space effective mass model with two-body carrier interactions. The interactions were taken to be of the Keldysh type, with screening lengths and effective masses taken from previous DFT parameterizations. Schrödinger's equation was solved exactly via diffusion Monte Carlo to obtain the ground state energies and radial distribution functions for the exciton, trion, and biexciton complexes. The results were compared with previous variational calculations and were found to be in quantitative agreement with the most accurate experimental estimates. Future work could apply DMC to study a gas of indirect excitons in TMDC heterostructures, which have been speculated to possibly exhibit high-temperature superfluidity.

CHAPTER 3

DOPING-DEPENDENT EXCITON AND TRION LINEWIDTHS

3.1 Introduction

In addition to exhibiting tightly bound excitonic complexes that allow for observation at high temperatures, two-dimensional transition metal dichalcogenides exhibit tunable optical properties via gating, which effectively introduces a finite density of electrons or holes into the material. Experiments performed on these materials have revealed that for doping densities corresponding to $\epsilon_{\mathrm{F}} \gtrsim \epsilon_{\mathrm{T}}$, where ϵ_{T} is the trion binding energy, the exciton peak in the optical absorption spectra is resolved as two separate peaks. The peaks are most commonly interpreted as arising from separate exciton and trion states. The splitting between these peaks and the linewidth of the excitonic peak exhibit a strong dependence on the splitting, sometimes increasing by dozens of meV over experimentally achievable doping ranges. In addition, very recent experiments performed on high quality samples at low temperatures place the intrinsic linewidth of the trion line in the range of several meV [100]. It is the purpose of this short chapter to propose a theory that explains this intrinsic linewidth.

In the previous chapter, a simple few-particle picture was developed and used along with the Keldysh form of the Coulomb interaction to describe exciton, trion, and biexciton binding energies in 2D transition metal dichalcogenides. The model almost perfectly reproduces all experimentally measured energies and validates the accuracy of the 2D Keldysh screening model for describing interactions between charged carriers. Here, we push the boundaries of our understanding of these materials as we try to describe the effects of interaction between these charge carrier complexes and a sea of free electrons.

We propose elastic scattering between charge carrier complexes and the free electrons introduced via doping as the dominant contribution to the linewidth broadening measured in experi-

ments. The results are calculated within first-order time-dependent perturbation theory and utilize both the Keldysh model for screening in 2D materials and the random phase approximation (RPA) model for screening due to a 2D electron gas.

3.2 Methods

3.2.1 Fermi's Golden Rule

Our procedure follows the one first reported by Cohen *et al.* [101]. We compute the elastic scattering rate within first-order time-dependent perturbation theory, i.e. Fermi's Golden Rule. Consider an elastic scattering process in which an electron of momentum $\boldsymbol{k}_e$ exchanges momentum $\boldsymbol{q}$ with an exciton (trion) of initial momentum $\boldsymbol{k}_X = 0$ ($\boldsymbol{k}_{X^-} = 0$) so that the final electron momentum is $\boldsymbol{k}_e - \boldsymbol{q}$ and the final exciton (trion) momentum is $\boldsymbol{q}$. The linewidth broadening due to the elastic version of this scattering process is then given by

$$\begin{aligned}\Gamma &= \frac{A^2}{4\pi^3}\int d^2\boldsymbol{k}_e \int d^2\boldsymbol{q}|V(\boldsymbol{k}_e,\boldsymbol{q})|^2 f(k_e)(1-f(|\boldsymbol{k}_e-\boldsymbol{q}|)) \\ &\quad \times\delta\left[E_e(k_e)+E_{X/X^-}(0)-(E_e(|\boldsymbol{k}_e-\boldsymbol{q}|)+E_{X/X^-}(q))\right],\end{aligned} \tag{3.1}$$

where A is the domain area which will later cancel, $V(\boldsymbol{k}_e,\boldsymbol{q})$ is the interaction matrix element for the scattering process, f is the Fermi-Dirac distribution function for the electron, $E_e(k) = \hbar^2k^2/2m_e$ is the kinetic energy of the electron, and $E_X(k) = \hbar^2k^2/2M_X$ ($E_{X^-}(k) = \hbar^2k^2/2M_{X^-}$) is the kinetic energy of the exciton (trion).

3.2.2 Exciton-electron scattering matrix elements

To compute $V(\boldsymbol{k}_e,\boldsymbol{q})$ for the exciton-electron scattering process, we first write the exciton wavefunction in terms of creation operators in momentum space. The real-space exciton wavefunction is assumed to be of Hydrogen 1s type:

$$\phi(r;\lambda) = \sqrt{\frac{2}{\pi\lambda^2}}e^{-r/\lambda}. \tag{3.2}$$

Now, the exciton state with momentum $\boldsymbol{k}_X$ can be written in terms of electron and hole creation operators $c^\dagger$ and $d^\dagger$ as

$$|\boldsymbol{k}_X\rangle = \sum_{\boldsymbol{k}_1} \phi^*_{\alpha \boldsymbol{k}_X + \boldsymbol{k}_1} c^\dagger_{-\boldsymbol{k}_1} d^\dagger_{\boldsymbol{k}_X + \boldsymbol{k}_1} |0\rangle, \tag{3.3}$$

where $\alpha = m_e/M_X$ and the Fourier-transformed exciton wavefunction is

$$\phi_{\boldsymbol{p}}(\lambda) = \frac{\sqrt{8\pi\lambda^2}}{(1+(\lambda p)^2)^{3/2}}. \tag{3.4}$$

The state of an exciton with momentum $\boldsymbol{k}_X$ and a free electron of momentum $\boldsymbol{k}_e$ can then be written as

$$|\boldsymbol{k}_X; \boldsymbol{k}_e\rangle = \sum_{\boldsymbol{k}_1} \phi^*_{\alpha \boldsymbol{k}_X + \boldsymbol{k}_1} \psi^*_{\boldsymbol{k}_e} c^\dagger_{-\boldsymbol{k}_1} d^\dagger_{\boldsymbol{k}_X + \boldsymbol{k}_1} c^\dagger_{\boldsymbol{k}_e} |0\rangle, \tag{3.5}$$

where $\psi_{\boldsymbol{k}_e}$ represents the free electron wavefunction.

The scattering matrix element can finally be written:

$$\langle \boldsymbol{k}_X + \boldsymbol{q}; \boldsymbol{k}_e - \boldsymbol{q} | V(\boldsymbol{q}) | \boldsymbol{k}_X; \boldsymbol{k}_e \rangle = \langle \boldsymbol{k}_X + \boldsymbol{q}; \boldsymbol{k}_e - \boldsymbol{q} | V_{ee}(\boldsymbol{q}) + V_{eh}(\boldsymbol{q}) | \boldsymbol{k}_X; \boldsymbol{k}_e \rangle, \tag{3.6}$$

where, for instance,

$$V_{ee}(\boldsymbol{q}) = \frac{2\pi e^2}{\epsilon_0 \epsilon(q) q} \sum_{\boldsymbol{k}_1, \boldsymbol{k}_2} c^\dagger_{\boldsymbol{k}_1 - \boldsymbol{q}} c^\dagger_{\boldsymbol{k}_2 + \boldsymbol{q}} c_{\boldsymbol{k}_2} c_{\boldsymbol{k}_1}. \tag{3.7}$$

Applying the fermionic anticommutation relations reveals

$$\langle \boldsymbol{k}_X + \boldsymbol{q}; \boldsymbol{k}_e - \boldsymbol{q} | V(\boldsymbol{q}) | \boldsymbol{k}_X; \boldsymbol{k}_e \rangle = V_{\text{dir}}(q) + V_{\text{exc}}(\boldsymbol{q}, \Delta \boldsymbol{k}_X), \tag{3.8}$$

where

$$V_{\text{dir}}(q) = \frac{2\pi e^2}{\epsilon_0 \epsilon(q) q A} \left[g(\lambda \beta a/2) - g(\lambda \alpha q/2) \right], \tag{3.9}$$

$$V_{\text{exc}}(\Delta \boldsymbol{k}_X, \boldsymbol{q}) = -\frac{4e^2\lambda^2}{\epsilon_0 A} \int d^2\boldsymbol{k}_1 \frac{1}{\epsilon(\boldsymbol{k}_1) k_1^3} g(\lambda |\boldsymbol{k}_1 + \alpha \boldsymbol{q} - \Delta \boldsymbol{k}|) \times \left[g(\lambda |\boldsymbol{k}_1 + \boldsymbol{q} - \Delta \boldsymbol{k}|) - g(\lambda |\boldsymbol{q} - \Delta \boldsymbol{k}|) \right], \tag{3.10}$$

where $\beta = 1 - \alpha$, $\Delta \boldsymbol{k}_X = \boldsymbol{k}_e - \alpha \boldsymbol{k}_X$, and

$$g(x) = \frac{1}{(1 + x^2)^{3/2}}. \tag{3.11}$$

3.2.3 Trion-electron scattering matrix elements

The calculation of the trion matrix elements is conceptually similar, but more complicated owing to the importance of electron spin. We take the singlet trion wavefunction to be of the Chandrasekhar type [102],

$$\phi^{X-}(r_{1\text{h}}, r_{2\text{h}}) = N_{X-} \left[\phi(r_{1\text{h}}; \lambda_1) \phi(r_{2\text{h}}; \lambda_2) + \phi(r_{1\text{h}}; \lambda_2) \phi(r_{2\text{h}}; \lambda_1) \right], \tag{3.12}$$

where $N_{X-} = 1/\sqrt{1 + \kappa^2}$ is the trion wavefunction normalization factor and $\kappa = 4\lambda_1\lambda_2/(\lambda_1 + \lambda_2)^2$.

Now, the trion state with momentum $\boldsymbol{k}_{X-}$ can be written in terms of electron and hole creation operators $c^\dagger$ and $d^\dagger$ as

$$|\boldsymbol{k}_{X-}\rangle = \sum_{\boldsymbol{k}_1, \boldsymbol{k}_2, s_1, s_2} \phi^{X-*}_{\alpha \boldsymbol{k}_{X-} + \boldsymbol{k}_1, \alpha \boldsymbol{k}_{X-} + \boldsymbol{k}_2} \xi^*(s_1, s_2) c^{s_1\dagger}_{-\boldsymbol{k}_1} c^{s_2\dagger}_{-\boldsymbol{k}_2} d^\dagger_{\boldsymbol{k}_t + \boldsymbol{k}_1 + \boldsymbol{k}_2} |0\rangle, \tag{3.13}$$

where $\alpha = m_e/M_{X-}$, $\xi(s_1, s_2) = \langle S | s_1, s_2 \rangle$ is the projection of the spin configuration $|s_1, s_2\rangle$ onto the singlet spin configuration, and the Fourier-transformed trion wavefunction is

$$\phi^{X-}_{\boldsymbol{p}_1, \boldsymbol{p}_2} = N_{X-} \phi_{\boldsymbol{p}_1}(\lambda_1) \phi_{\boldsymbol{p}_2}(\lambda_2). \tag{3.14}$$

The state of a trion with momentum $\boldsymbol{k}_{X-}$ and a free electron of spin s_e and momentum $\boldsymbol{k}_e$ can then be written as

$$|\boldsymbol{k}_{X-};\boldsymbol{k}_e\rangle = \sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,s_1,s_2,s_e} \phi^{X-*}_{\alpha\boldsymbol{k}_{X-}+\boldsymbol{k}_1,\alpha\boldsymbol{k}_{X-}+\boldsymbol{k}_2}\psi^*_{\boldsymbol{k}_e}\xi^*(s_1,s_2)c^{s_1\dagger}_{-\boldsymbol{k}_1}c^{s_2\dagger}_{-\boldsymbol{k}_2}d^\dagger_{\boldsymbol{k}_t+\boldsymbol{k}_1+\boldsymbol{k}_2}c^{s_e\dagger}_{\boldsymbol{k}_e}|0\rangle. \tag{3.15}$$

The scattering matrix element can be written as

$$\langle \boldsymbol{k}_{X-}+\boldsymbol{q};\boldsymbol{k}_e-\boldsymbol{q}|V(\boldsymbol{q})|\boldsymbol{k}_{X-};\boldsymbol{k}_e\rangle = V_{\text{dir}}(q)+V^{(1)}_{\text{exc}}(\boldsymbol{q},\Delta\boldsymbol{k}_{X-})+V^{(2)}_{\text{exc}}(\boldsymbol{q},\Delta\boldsymbol{k}_{X-}), \tag{3.16}$$

where

$$\begin{aligned}
V_{\text{dir}}(q) &= N_{\text{es}}\{[g(\lambda_1\alpha q/2)g(\lambda_2\beta q/2)+g(\lambda_2\alpha q/2)g(\lambda_1\beta q/2) \\
&\quad +2\kappa^2 g(\tilde{\lambda}\alpha q)g(\tilde{\lambda}\beta q)] \\
&\quad -[g(\lambda_1\alpha q/2)g(\lambda_2\alpha q/2)+\kappa^2 g^2(\tilde{\lambda}\alpha q)]\},
\end{aligned} \tag{3.17}$$

$$\begin{aligned}
V^{(1)}_{\text{exc}}(\boldsymbol{q},\Delta\boldsymbol{k}_{X-}) &= -\frac{N_{\text{es}}}{\pi}\int d^2\boldsymbol{k}_1\{[g(\lambda_2|\boldsymbol{k}_1-\Delta\boldsymbol{k}_{X-}+\boldsymbol{q}|)-g(\lambda_2|\Delta\boldsymbol{k}_{X-}-\boldsymbol{q}|)] \\
&\quad \times[\lambda_1\lambda_2\kappa g(\tilde{\lambda}\alpha q)g(\lambda|\boldsymbol{k}_1-\Delta\boldsymbol{k}_{X-}+\alpha\boldsymbol{q}|) \\
&\quad +\lambda_2^2 g(\lambda_1\alpha q/2)g(\lambda_2|\boldsymbol{k}_1-\Delta\boldsymbol{k}_{X-}+\alpha\boldsymbol{q}|)] \\
&\quad +[\lambda_1\leftrightarrow\lambda_2]\times[\lambda_1\leftrightarrow\lambda_2]\},
\end{aligned} \tag{3.18}$$

$$\begin{aligned}
V^{(2)}_{\text{exc}}(\boldsymbol{q},\Delta\boldsymbol{k}_{X-}) &= -\frac{N_{\text{es}}}{\pi}\int d^2\boldsymbol{k}_1\{g(\lambda_1|\boldsymbol{k}_1-\Delta\boldsymbol{k}_{X-}+\alpha\boldsymbol{q}|) \\
&\quad \times[\lambda_1\lambda_2\kappa g(\tilde{\lambda}|\boldsymbol{k}_1-\alpha\boldsymbol{q}|)g(\lambda_2|\boldsymbol{q}-\Delta\boldsymbol{k}_{X-}|) \\
&\quad +\lambda_1^2 g(\lambda_2|\boldsymbol{k}_1-\alpha\boldsymbol{q}|/2)g(\lambda_1|\boldsymbol{q}-\Delta\boldsymbol{k}_{X-}|)] \\
&\quad +[\lambda_1\leftrightarrow\lambda_2]\times[\lambda_1\leftrightarrow\lambda_2]\},
\end{aligned} \tag{3.19}$$

where $\beta = 1-\alpha$, $\Delta\boldsymbol{k}_{X-} = \boldsymbol{k}_e - \alpha\boldsymbol{k}_{X-}$, $\tilde{\lambda} = \lambda_1\lambda_1/(\lambda_1+\lambda_2)^2$, and

$$N_{\text{es}} = \frac{1}{1+\kappa^2}\frac{4\pi e^2\lambda_1^3}{\epsilon_0}. \tag{3.20}$$

3.2.4 Material parameters for MoS_2

The relevant material parameters for monolayer MoS_2 are $\epsilon_\infty = 5.2$, $\lambda_1 = 10$ Å, $\lambda_2 = 25$ Å, and $m_e = m_h = 0.32m_0$, where m_0 is the rest electron mass. We use an electron density-dependent screening model derived from the random-phase approximation (RPA, or sum of "bubble diagrams") in two dimensions,

$$\epsilon(q) = \begin{cases} 1 + k_{\text{TF}}/q & q \leq 2k_{\text{F}} \\ 1 + \dfrac{k_{\text{TF}}}{q}\left[1 - \sqrt{1 - (2k_{\text{F}}/q)^2}\right] & q > 2k_{\text{F}}. \end{cases}$$

Unlike the Thomas-Fermi screening model, this model gives $\epsilon = 1$ in the $n \to 0$ (unscreened) limit.

For the degenerate 2D electron gas we are considering, we have $k_{\text{F}} = \sqrt{2\pi n}$, where n is the electron density. The chemical potential is $\mu = k_{\text{B}}T\ln(e^{E_{\text{F}}/k_{\text{B}}T} - 1)$ and the Thomas-Fermi screening wavevector is $k_{\text{TF}} = 2m_e e^2/\hbar^2$.

3.3 Results

In Fig. 3.1, we present the exciton linewidth broadening due to elastic electron-exciton scattering for various doping densities. Experiments done for this work measure little linewidth boradening increase from $n = 0$ to 4×10^{12} cm^{-2} doping, and then an approximately linear increase of about 40 meV in linewidth broadening from $n = 4 \times 10^{12}$ cm^{-2} to 8×10^{12} cm^{-2} at both $T = 50$ K and $T = 290$ K.

Clearly, the calculated linewidth broadening due to elastic scattering deviates largely from the experimental result. The decrease in the calculated linewidth broadening after peaks at $n = 10^{12}$ cm^{-2} and $n = 3 \times 10^{12}$ cm^{-2} at $T = 50K$ and $T = 290K$, respectively, is due to the $1 - f_e$ factor in Eq. 3.1. This effect, called Fermi blocking, leads to low scattering rates at high doping densities because simultaneous energy and momentum conservation becomes difficult to satisfy when the electron momentum is large and close to the Fermi surface.

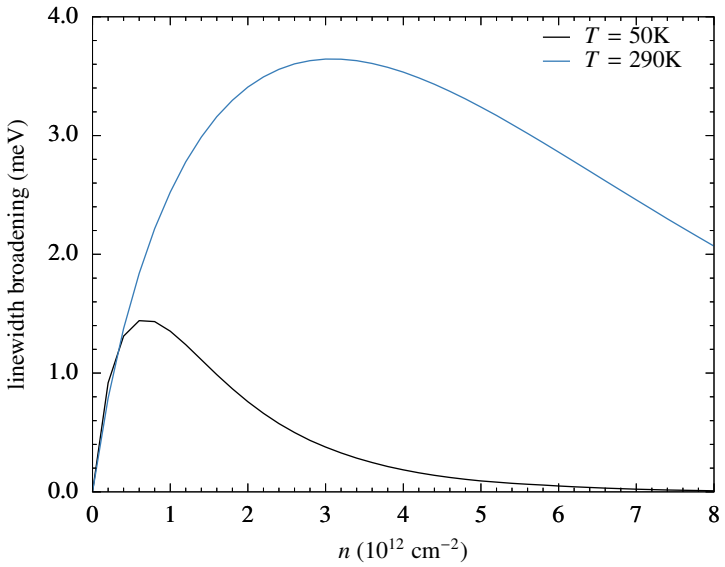

Figure 3.1: The doping-dependent linewidth broadening due to elastic electron-exciton scattering at $T = 50$ K and $T = 290$ K.

In Fig. 3.2, the calculated trion linewidth broadening due to elastic trion-electron scattering is plotted along with the experimentally-measured intrinsic linewidth. The Fermi blocking effect is observed to be less significant in the trion case as the linewidth broadening remains fairly constant for $n = 0$ to 2×10^{12} cm^{-2} before increasing nearly linearly from $n = 2 \times 10^{12}$ cm^{-2} to 8×10^{12}

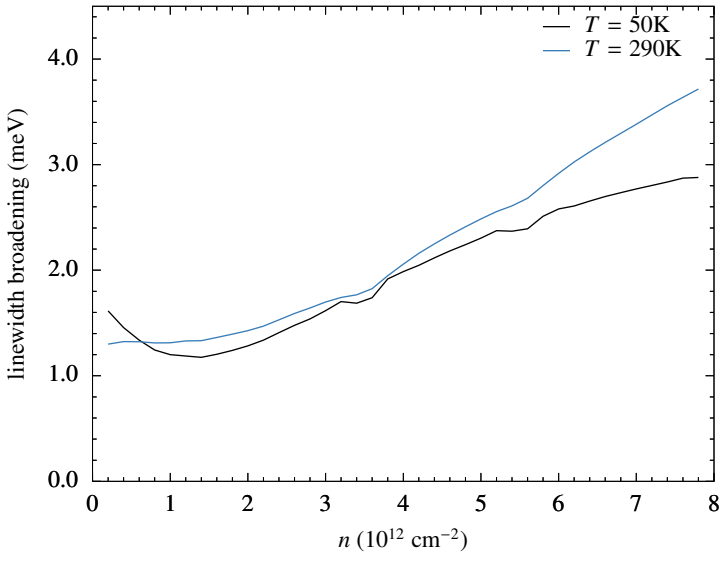

Figure 3.2: The doping-dependent linewidth broadening due to elastic electron-trion scattering at $T = 50$ K and $T = 290$ K.

cm^{-2}. These results are extremely similar to the experimental measurements, for which only small changes of about $2-5$ meV are observed throughout the measured doping range.

This is the first study to calculate the intrinsic trion linewidth due to elastic scattering. Past studies which have predicted the nearly linear increase in electron-exciton scattering rate as well as the constant electron-trion scattering rate have always used an adjustable parameter to fit the (predicted constant) trion linewidth. As mentioned in the introduction to this section, recent experiments place the intrinsic trion linewidth at approximately $1-4$ meV [100]. Quite encouragingly, our results from Fig. 3.2 are in quantitative agreement.

3.4 Conclusions

In this work electron and trion linewidths, as have been measured experimentally in WS_2 as a function of the carrier doping density at two different temperatures [103] were studied. At both temperatures, a large exciton linewidth increase of about 40 meV was measured over the studied doping range, while only modest increases of about 2-5 meV in the trion linewidth were measured. These experimental observations were proposed to arise from an elastic scattering process between the exciton or trion and the free electrons present due to the doping of the material. The elastic scattering model reproduces the trion linewidth results well (including the magnitude of the intrinsic width of ~ 1 meV as revealed by very recent experiments [100]) and provides the first calculation of the intrinsic trion linewidth. However, the model fails to reproduce the large increase in the exciton linewidth at large doping densities due to the Fermi blocking phenomenon. This shortcoming will be resolved in the published work using a non-perturbative approach [104].

CHAPTER 4
A NEW GW CUMULANT APPROXIMATION

4.1 Introduction

The development of methods that can accurately and affordably describe both the total electronic energy and the electronic excitations of complex systems remains a long-standing challenge in both condensed matter physics and chemistry. Substantial progress has been made in recent decades on the development of approximate approaches to calculate correlation contributions that go beyond the Hartree-Fock level of mean-field theory. While density functional theory (DFT), including its extensions to hybrid functionals, has emerged as an accurate and efficient means of calculating many properties of both solids and molecules, systematic improvement of DFT is challenging [105]. In particular, while the Kohn-Sham eigenvalues in the theory often give a remarkably accurate band structure, there remain fundamental shortcomings with respect to properly calculated excitation energies [106, 105]. More broadly, considering the proliferation of a myriad of approximate exchange-correlation functionals, care must be taken in applications to assess empirical evidence of the accuracy for specific classes of materials.

Separate from DFT, direct many-body methods based on wavefunctions have achieved impressive accuracy, exemplified by coupled-cluster methods for finite systems [107] and quantum Monte Carlo (QMC) methods for extended systems [80]. These approaches are more numerically challenging than DFT due to unfavorable scaling with system size (or complexity), but they are often regarded as a "gold standard" when they can be applied. They are also typically more difficult to apply to excited state properties with the same accuracy. Nonetheless, recent progress is encouraging for more widespread application to solids [108, 109, 110].

On the other hand, the more general language of Green's function-based perturbation expansions allows for the natural description of spectral features and quasiparticle properties of extended

systems in a size-consistent manner [111]. In particular, the development of the GW approximation for calculation of the one-particle Green's function [112] led to the first predictive calculations of charged excitation energies in real materials [113, 114, 115, 116]. The extension to the Bethe-Salpeter equation the two-particle Green's function has correspondingly supported calculation of the neutral excitations, such as those probed in optical absorption [106]. The development of systematic corrections beyond GW, including approximations to the vertex corrections and the use of self-consistency, remains a subject of ongoing research [117, 118, 119, 120, 121, 122, 123, 124, 125]. Interestingly, the corresponding Green's function-based method for the total electronic energy has not been widely used, although several formulations have been investigated [119, 126, 127].

The homogeneous electron gas model in three dimensions (3D) has been widely used as a model system for high-level electronic structure calculations because it captures essential features of the electronic structure of simple metals while retaining a higher symmetry that allows for some integrations to be performed analytically. In spite of this higher symmetry, no exact solution for the Green's function for this system exists, although practically exact correlation energies and momentum-dependent occupation numbers have been calculated using different quantum Monte Carlo variations (including the diffusion Monte Carlo method described in Chapter 2) [128]. Results based on G_0W_0 (the non-self-consistent first iteration of the GW approximation) show very reasonable quasiparticle properties, but also an unphysical satellite structure ("plasmaron peak") about $1.5\omega_p$ below the quasiparticle peak, where ω_p is the plasmon energy [129]. This is a surprising result since standard coupled electron-boson models would suggest a series of satellite peaks near integer multiples of ω_p below the quasiparticle peaks [21]. Calculations in which G was iterated to full self-consistency, conceptually part of Hedin's original framework [112], indicated further unphysical changes in the satellite region [117, 118], although the accuracy of such calculations has recently been called into doubt [125]. Self-consistent GW gives reasonable correlation energies, but it was suggested that either vertex corrections [126] or an alteration to the calculated polarization required to conserve particle number [125] were needed in addition to restore physical

spectral properties.

The difficulty of describing satellite structures in the spectral function strongly suggests a cumulant-based approach to approximately include vertex corrections [130]. This idea has been extensively explored with a time-ordered formulation of G [131, 132, 133, 134, 135, 123, 136], and has been successful in describing the satellite structure in metals and semiconductors [137, 138, 139, 140, 141]. The approach results in the expected satellite progression and slightly modifies the quasiparticle properties. In part, the exponential form imposed by the cumulant ansatz leads to the inclusion of higher-order exchange-like diagrams that are only accessible in the standard GW formalism by way of vertex corrections. However, these higher-order diagrams do *not* correspond exactly to standard diagrams in the time-ordered Dyson expansion. More generally, the cumulant approach has not yet reached the formal level of sophistication that is afforded by the standard diagrammatic apparatus that surrounds the Dyson equation. In particular, aspects related to self-consistency, conservation laws, and the one-to-one correspondence of terms within the cumulant expansion to standard Feynman-Dyson diagrammatics require further investigation.

The time-ordered cumulant approach is limited by the serious drawback that it precludes the possibility of positive spectral weight both above and below the chemical potential. Recently, Kas *et al.* showed that the retarded Green's function is a more natural quantity to employ with the cumulant formalism, as it allows for the description of both particles and holes within one spectral weight profile [142]. Compared with standard G_0W_0 [129], the retarded cumulant approach predicts more physical satellite properties, similar energies, but somewhat less accurate wavevector-dependent occupation numbers for the 3D electron gas.

In this chapter, we take this retarded Green's function perspective as a starting point to investigate and compare new cumulant-based GW schemes. We begin by giving an overview of the homogeneous electron gas model before reviewing the Green's function formalism and defining the G_0W_0 approximation and related cumulant methods. Each approach is applied to calculate the Green's function for a system consisting of a single electronic state coupled to independent bosons, for which the exact solution is derived in an appendix. The methods are then applied to the

calculation of the electronic Green's function for the full electron gas model system. Wavevector-dependent occupation numbers and correlation energies are also calculated and compared to those that result from standard G_0W_0 calculations and the known exact answers.

4.2 The electron gas (jellium) model

The 3-dimensional electron gas consisting of N particles in a volume Ω is described by the Hamiltonian

$$\begin{aligned} H &= H_- + H_{-+} + H_{++}, && (4.1)\\ H_- &= \sum_{i=1}^{N} \frac{-\hbar^2}{2m_e}\nabla_i^2 + \sum_{i<j}^{N} e^2 V(|\boldsymbol{r}_i - \boldsymbol{r}_j||), && (4.2)\\ H_{-+} &= -\sum_{i=1}^{N} \int_\Omega d\boldsymbol{R} n e^2 V(|\boldsymbol{R} - \boldsymbol{r}_i|), && (4.3)\\ H_{++} &= \int_\Omega d\boldsymbol{R}_1 \int_\Omega d\boldsymbol{R}_2 n^2 \frac{e^2}{2} V(|\boldsymbol{R}_1 - \boldsymbol{R}_2|), && (4.4) \end{aligned}$$

where $V(r) = 1/r$ is the Coulomb potential and Ω represents the volume of the system, which is usually taken to be cubic and periodic with some constant electron density $n = N/\Omega$. After careful determination of the action of the electronic creation and destruction operators upon a wavefunction in the many-particle Hilbert space, the electronic part of the Hamiltonian can be written using second quantization as

$$H_- = \sum_{rs} \langle r|T|s\rangle c_r^\dagger c_s + \frac{1}{2}\sum_{rstu} \langle rs|V|tu\rangle c_r^\dagger c_s^\dagger c_u c_t, \tag{4.5}$$

where the indices run over all one-electron states with the right symmetry. For the electron gas, these states are $\psi_{\boldsymbol{k}\lambda}(\boldsymbol{r}) = \frac{1}{\Omega^{1/2}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\eta_\lambda$, where $\boldsymbol{k} = \frac{2\pi}{a}(k_x, k_y, k_z)$ for k_x, k_y, k_z integers, $\Omega = a^3$, and η_λ is a spin function. The electronic creation and destruction operators are defined by

$$\{c_r, c_s\} = \{c_r^\dagger, c_s^\dagger\} = 0, \tag{4.6}$$

$$\{c_r, c_s^\dagger\} = \delta_{rs}, \tag{4.7}$$

where the anticommutator $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} - \hat{B}\hat{A}$.

With these wavefunctions, the matrix elements in Eq. 4.5 are readily evaluated. The second-quantized Hamiltonian becomes

$$H_- = \sum_{\boldsymbol{k}\lambda} \frac{\hbar^2 k^2}{2m} c_{\boldsymbol{k}\lambda}^\dagger c_{\boldsymbol{k}\lambda} + \frac{e^2}{2\Omega} \sum_{\substack{\boldsymbol{kpq} \\ \lambda_1\lambda_2}} \frac{4\pi}{q^2} c_{\boldsymbol{k}+\boldsymbol{q},\lambda 1}^\dagger c_{\boldsymbol{p}-\boldsymbol{q},\lambda_2}^\dagger c_{\boldsymbol{p}\lambda_2} c_{\boldsymbol{k}\lambda_1}. \tag{4.8}$$

The $\boldsymbol{q} = 0$ term of the latter sum cancels exactly with $H_{+-} + H_{++}$ so that the overall electron gas Hamiltonian becomes

$$H = \sum_{\boldsymbol{k}\lambda} \frac{\hbar^2 k^2}{2m} c_{\boldsymbol{k}\lambda}^\dagger c_{\boldsymbol{k}\lambda} + \frac{e^2}{2\Omega} \sum_{\substack{\boldsymbol{kp} \\ \boldsymbol{q}\neq 0 \\ \lambda_1\lambda_2}} \frac{4\pi}{q^2} c_{\boldsymbol{k}+\boldsymbol{q},\lambda 1}^\dagger c_{\boldsymbol{p}-\boldsymbol{q},\lambda_2}^\dagger c_{\boldsymbol{p}\lambda_2} c_{\boldsymbol{k}\lambda_1}. \tag{4.9}$$

It is customary to write this simplified Hamiltonian in terms of dimensionless variables $N/\Omega = (\frac{4}{3}\pi(r_s a_0)^3)^{-1}$, where $a_0 = \hbar^2/me^2$ is the Bohr radius. The Wigner-Seitz radius r_s is a measure of the system's density; many model metals of interest, including Sodium, have $1 \leq r_s \leq 5$. After making the substitutions $\overline{\Omega} = (r_s a_0)^{-3}\Omega$, $\overline{\boldsymbol{k}} = (r_s a_0)\boldsymbol{k}$, $\overline{\boldsymbol{p}} = (r_s a_0)\boldsymbol{p}$, $\overline{\boldsymbol{q}} = (r_s a_0)\boldsymbol{q}$, the system Hamiltonian can be written as

$$H = \frac{e^2}{a_0 r_s^2} \left(\sum_{\overline{\boldsymbol{k}}\lambda} \frac{\overline{k}^2}{2} c_{\overline{\boldsymbol{k}}\lambda}^\dagger c_{\overline{\boldsymbol{k}}\lambda} + \frac{r_s}{2\overline{\Omega}} \sum_{\substack{\overline{\boldsymbol{k}}\overline{\boldsymbol{p}} \\ \overline{\boldsymbol{q}}\neq 0 \\ \lambda_1\lambda_2}} \frac{4\pi}{\overline{q}^2} c_{\overline{\boldsymbol{k}}+\overline{\boldsymbol{q}},\lambda_1}^\dagger c_{\overline{\boldsymbol{p}}-\overline{\boldsymbol{q}},\lambda_2}^\dagger c_{\overline{\boldsymbol{p}}\lambda_2} c_{\overline{\boldsymbol{k}}\lambda_1} \right), \tag{4.10}$$

which makes evident the fact that the potential energy is a small perturbation in the limit $r_s \to 0$. This is quite astonishing, as the $r_s \to 0$ limit is the limit of large electron density, i.e. *more*

electrons!

For small r_s, the ground-state energy can be approximated by solving the non-interacting (kinetic) part exactly, and then adding in the first-order contribution due to the interactions. Since the non-interacting part is diagonal, its ground-state wavefunction will be the direct product of the N lowest-energy plane waves. Using this state, the analytic form for both the non-interacting energy and the first-order correction to the total energy can be determined an a straightforward manner.

The full ground-state energy for this model takes the form [143]

$$E = \frac{2Ne^2}{a_0 r_s^2}(a + br_s + cr_s^2 \ln r_s + dr_s^2 + \ldots), \tag{4.11}$$

where $a = \frac{3}{5}\left(\frac{9\pi}{4}\right)^{2/3} \approx 2.21$ Rydbergs and $b = -\frac{3}{2\pi}\left(\frac{9\pi}{4}\right)^{1/3} \approx -0.916$ Rydberg. If this series is truncated after the first two terms, the result is the Hartree-Fock energy.

4.3 The Green's function

The single-particle Green's function is defined as

$$iG_{\alpha\beta}(\boldsymbol{x}t, \boldsymbol{x}'t') = \frac{\langle\Psi_0|T[\hat{\psi}_{H\alpha}(\boldsymbol{x}t)\hat{\psi}^\dagger_{H\beta}(\boldsymbol{x}'t')]|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle}, \tag{4.12}$$

where Ψ_0 is the ground state of the Hamiltonian H, $\hat{\psi}_{H\alpha}(\boldsymbol{x}t) = e^{i\hat{H}t/\hbar}\hat{\psi}_\alpha(\boldsymbol{x})e^{-i\hat{H}t/\hbar}$, T is the time-ordering operator which places operators with later time arguments to the left of those with earlier time arguments, and $\hat{\psi}$ is the field operator, defined by

$$\hat{\psi}(\boldsymbol{x}) = \sum_{k\alpha} \psi_{k\alpha}(\boldsymbol{x})c_k, \tag{4.13}$$

where the ψ_k run over all single-particle wavefunctions, which for the electron gas are the plane waves defined above.

The Green's function is a useful quantity because it can be used to determine the expected value of all one-particle operators. It can also be used to obtain the energy of the ground-state,

which is comprised of the one-particle operator $\hat{T}$ and the two-particle operator $\hat{V}$. Physically, the Green's function describes the overlap between $\hat{U}(t,t')\hat{\psi}^{\dagger}_{I\beta}(\boldsymbol{x}'t')|\Psi_I(\boldsymbol{t}')\rangle$ and $\hat{\psi}^{\dagger}_{I\beta}(\boldsymbol{x}t)|\Psi_I(t)\rangle$, where $\hat{U}$ denotes the time-propagator and the index I specifies quantities in the interaction picture. These quantities correspond to a particle being created at $(\boldsymbol{x}', t')$ and propagated from t' to t (the former) and a particle being created at $(\boldsymbol{x}t)$ (the latter).

The Fourier transform of the non-interacting Green's function (i.e. the Green's function for the Hamiltonian without the interacting part) is

$$G^0_{\alpha\beta}(\boldsymbol{k},\omega) = \delta_{\alpha\beta}\left[\frac{\theta(k-k_F)}{\omega-\omega_k+i\eta}+\frac{\theta(k_F-k)}{\omega-\omega_k-i\eta}\right]. \tag{4.14}$$

In essence, this describes the particle accumulating a phase $e^{i\epsilon_k(t'-t)/\hbar}$ in the time $t'-t$, where $\epsilon_k = \hbar^2k^2/2m$. The infinitesimal parts result from the time-ordering operator in the Green's function. The θ-functions describe which states are filled in the ground state (those with $k < k_F$).

Now, the full Green's function can be written using a linked-cluster expansion as

$$iG_{\alpha\beta}(x,y) = \sum_{v=0}^{\infty}\left(\frac{-i}{\hbar}\right)^v\frac{1}{v!}\int_{-\infty}^{\infty}dt_1\ldots\int_{-\infty}^{\infty}dt_v\frac{\langle\Phi_0|T[\hat{H}_1(t_1)\ldots\hat{H}_1(t_v)\hat{\psi}_\alpha(x)\hat{\psi}^{\dagger}_\beta(y)]|\Phi_0\rangle}{\langle\Phi_0|\hat{U}(-\infty,\infty)|\Phi_0\rangle}, \tag{4.15}$$

where $|\Phi_0\rangle$ is the non-interacting ground state, $\hat{H}_1$ is the interacting part of the Hamiltonian, and $\hat{U}$ is the time-development operator. After writing the $\hat{H}_1$ in terms of field operators via

$$\int_{-\infty}^{\infty}dt\hat{H}_1(t)\to\sum_{\lambda\lambda'\mu\mu'}\int\ldots\int d^3\boldsymbol{x}_1d^3\boldsymbol{x}'_1dt_1dt'_1\hat{\psi}^{\dagger}_\lambda(\boldsymbol{x}_1)\hat{\psi}^{\dagger}_\mu(\boldsymbol{x}'_1)\hat{\psi}_{\mu'}(\boldsymbol{x}'_1)\hat{\psi}_{\lambda'}(\boldsymbol{x}_1)U(\boldsymbol{x}_1,\boldsymbol{x}'_1)\delta(t_1-t'_1), \tag{4.16}$$

each term of Eq. 4.15 can be evaluated using an application of Wick's theorem.

The idea behind Wick's theorem is to move all the destruction operators (which are present in the field operators) to the right of the creation operators. Upon each reordering of field operators, the fermionic commutation relations are applied and the resulting terms are carried along. In the end, the full Green's function G can be expressed as a sum of terms containing the interaction U

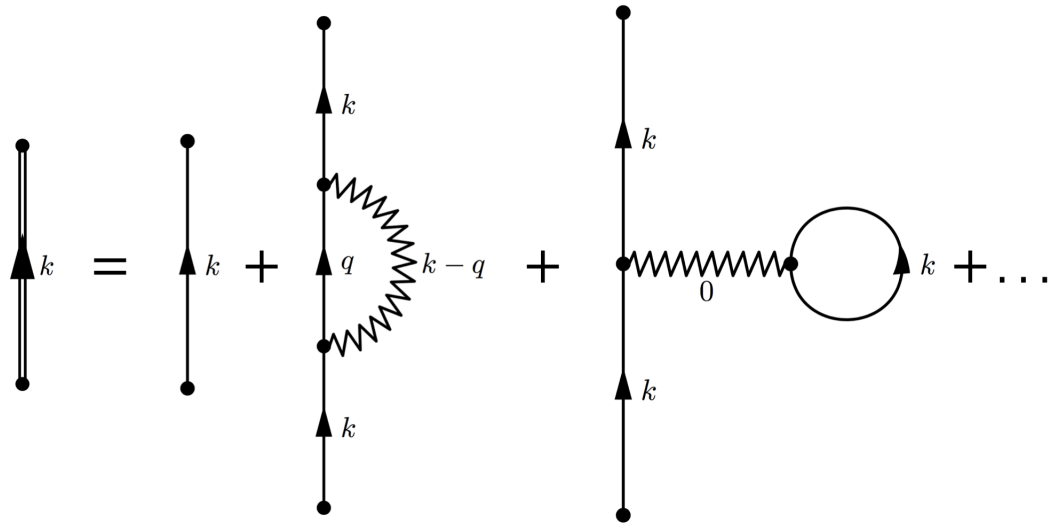

Figure 4.1: Feynman diagrams for the zeroth- and first-order terms in the Green's function expansion.

and the non-interacting Green's function G_0. The first few terms look like

$$\begin{aligned} G_k(\omega) &= G_k^0(\omega) + \frac{i}{\hbar(2\pi)^4}\int d^3\boldsymbol{q}d\omega' G_k^0(\omega)U_{|\boldsymbol{k}-\boldsymbol{q}|}G_q^0(\omega')G_k^0(\omega) \\ &\quad -\frac{i}{\hbar(2\pi)^4}\int d^3\boldsymbol{q}d\omega' G_k^0(\omega)U_0 G_k^0(\omega)G_q^0(\omega') + \ldots \end{aligned} \tag{4.17}$$

Each term of G can be depicted using a Feynman diagram, the simplest of which are drawn in Fig. 4.1.

Evaluation of the terms in the Green's function expansion quickly becomes difficult with increasing order, as each n^{th}-order term is a $4n$-dimensional integral. Fortunately, the expansion can be rewritten ("resummed") using various tricks so that a representative infinite subset of the terms *can* be evaluated fairly efficiently, even for material systems.

The many-body perturbation expansion for the one-particle Green's function is resummed via

the Dyson equation,

$$
\begin{aligned}
G_k(\omega) &= G_k^0(\omega) + G_k^0(\omega)\Sigma_k^I(\omega)G_k^0(\omega), && (4.18)\\
&= G_k^0(\omega)[1 + \Sigma_k^I(\omega)G_k^0(\omega)] && (4.19)\\
&= \frac{G_k^0(\omega)}{1 - \Sigma_k^*(\omega)G_k^0(\omega)}, && (4.20)
\end{aligned}
$$

where $G_k^0(\omega)$ is the non-interacting Green's function, $\Sigma_k^*(\omega)$ is the proper self-energy, and $\Sigma_k^I(\omega)$ is the improper self-energy [111]. The improper self-energy maps directly onto the Green's function expansion and contains every diagram without the leading and trailing $G_k^0(\omega)$ (cf. Eq. 4.19). The *proper* self-energy, on the other hand, contains only those diagrams which are not broken by a cutting a bare propagator (i.e. G_0) line. Use of the proper self-energy allows for infinite subsets of diagrams to be summed at the cost of explicitly evaluating only low-order ones. The proper self-energy is defined schematically in Fig. 4.2a.

If the proper self-energy is truncated at first-order, the resulting theory can be shown to be exactly Hartree-Fock. Although it is certainly a step in the right direction from the non-interacting solution, it hardly justifies the formalism that has so far been developed. In fact, a larger subset of diagrams can be summed at limited additional cost by resumming the interaction as well, leading to a screened interaction. Analogously to the self-energy, we define the improper and proper *polarization insertions* Π^I and Π^* to include those diagrams which are included in the full interaction expansion. The proper polarization insertion is defined schematically in Fig. 4.2b. When the proper polarization insertion is truncated at first-order (i.e. includes only the "bubble diagram"), the result is the random-phase approximation (RPA) for electronic screening. The RPA model for screening is widely used in electronic structure calculations of metals and semiconductors because it captures important qualitative features of the electron gas, most notably plasma oscillations.

We are now ready to examine the GW approximation. In the non-self-consistent GW approximation (G_0W_0), $\Sigma_k^*(\omega)$ is truncated at first order and the RPA $W_k(\omega)$ is used in place of the bare Coulomb interaction v_k [111]:

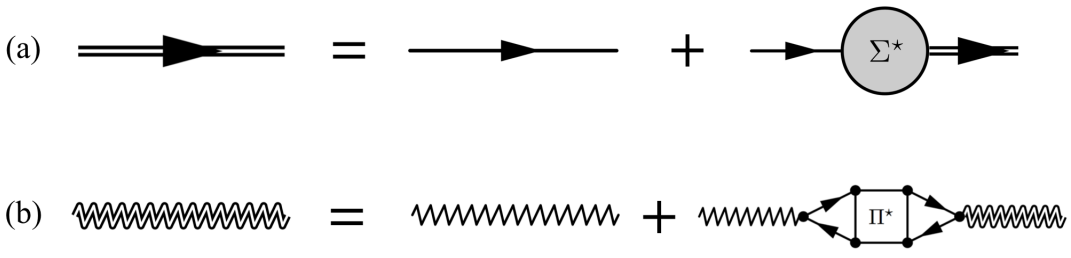

Figure 4.2: Definitions of the (a) proper self-energy and (b) proper polarization insertion.

$$\Sigma_k^*(\omega) = \frac{i}{\hbar}\frac{1}{(2\pi)^4}\int d^3q d\omega' G_k(\omega) W_{k-q}(\omega-\omega'), \tag{4.21}$$

$$W_k(\omega) = \frac{v_k}{1-\Pi_k^*(\omega)v_k}, \tag{4.22}$$

$$\Pi_k^*(\omega) = \frac{i}{\hbar}\frac{1}{(2\pi)^4}\int d^3q d\omega' G_q(\omega) G_{k+q}(\omega-\omega'). \tag{4.23}$$

These equations are summarized schematically in Fig. 4.3.

Lundqvist [129] was the first to numerically evaluate the G_0W_0 self-energy in Eq. 4.21 for the homogeneous electron gas. The azimuthal angle integration can be trivially completed. Next, the contour of the frequency integral is turned to the imaginary axis. Residues that are collected result in one two-dimensional integral; the remaining contour along the imaginary-ω axis results in a

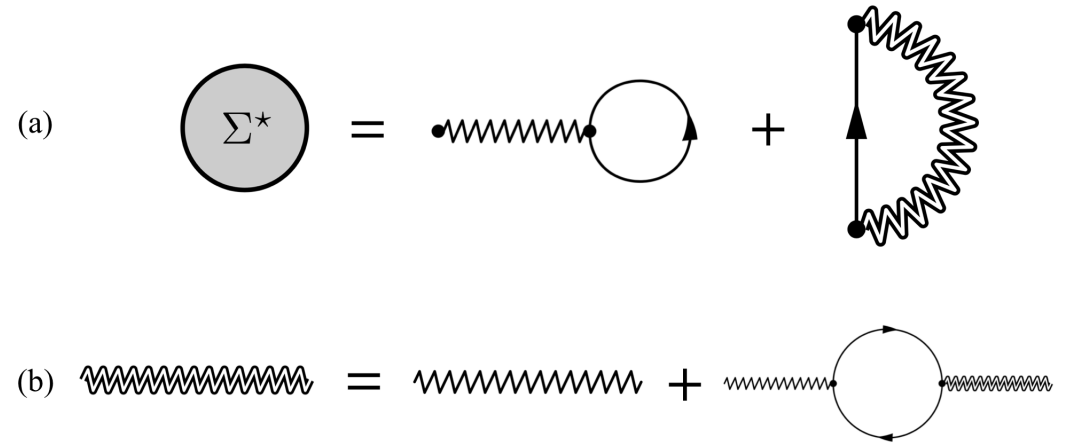

Figure 4.3: The (a) self-energy and (b) polarization insertion used in the non-self-consistent GW (or G_0W_0) approximation.

three-dimensional integral which can be analytically simplified to a two-dimensional integral. (Interestingly, the simplification of this latter integral cannot be done for the two-dimensional electron gas, i.e. it is *more* expensive to evaluate the self-energy for the system of lower dimensionality!)

The G_0W_0 approximation to the Green's function results in a dramatic improvement to density functional descriptions of electronic structure. In particular, G_0W_0 allows for finite lifetime of states and a renormalization of the quasiparticle peak. It predicts a fairly accurate band structure (cf. Fig. 4.5a below), momentum-dependent occupation numbers (Fig. 4.6), as well as correlation energies (Fig. 4.7 and Table 4.1). Perhaps the most noticeable qualitative deficiency of G_0W_0 is that it predicts an unphysical spacing between the quasiparticle peak and first satellite at the bottom of the band ($k \ll k_F$); plasma oscillations of frequency ω_{P} are the main source of scattering at the bottom of the band, yet G_0W_0 predicts a satellite approximately $1.5\omega_{\mathrm{P}}$ lower in energy than the quasiparticle peak. Historically, the state associated with this peak was given the name "plasmaron," but today it is generally viewed as a spurious feature of the G_0W_0 spectral weight.

Motivated by this discrepancy, a more recent approach to calculating the Green's function has been developed. The G_0W_0 plus cumulant method exactly reproduces the solution for a single electronic state coupled to independent phonons (cf. Eq. 4.28 below), and we will see that it also predicts a more physical spacing of the satellite peaks at the bottom of the band.

The retarded cumulant ansatz is a resummation of Eqs. 4.18. It can be written as [142]

$$G_k^R(t,T) = G_k^{0,R}(t,T)e^{C_k^R(t,T)}, \tag{4.24}$$

where $C_k(t,T)$ is the time-local cumulant function and the "R" superscripts denote retarded quantities.

When considered with Eq. 4.19, the cumulant ansatz for the retarded one-particle Green's function leads to a simple closed and exact relation between the *improper* Dyson self-energy and

the cumulant generating function:

$$C_k(t,t') = \ln\left(1 + \left[G_k^{R,0}(t,t')\right]^{-1} \int\int dt_1 dt_2 G_k^{R,0}(t,t_1)\Sigma_k^{R,I}(t_1,t_2)G_k^{R,0}(t_2,t')\right). \quad (4.25)$$

While it is clear that Eq. 4.25 trivially reduces to the standard Dyson equation, it should be noted that such a simple direct and exact relationship between the retarded cumulant and improper retarded Dyson self-energy has, to the best of our knowledge, not been noted before. Such a relation is only possible when retarded quantities and the improper as opposed to the proper self-energy are used. This relation implies new cumulant-like approximations distinct from earlier formulations. For example, the lowest order expansion of the logarithm in conjunction with a retarded, improper self-energy calculated using the normal first-order GW diagrams yields

$$C_k(t,t') = \left[G_k^{R,0}(t,t')\right]^{-1} \int\int dt_1 dt_2 G_k^{R,0}(t,t_1)\Sigma_{GW,k}^{R,I}(t_1,t_2)G_k^{R,0}(t_2,t'). \quad (4.26)$$

This equation for the cumulant (henceforth referred to as G_0W_0 with improper retarded cumulant, or G_0W_0 IRC) is superficially nearly identical to the cumulant approach of Kas *et al.* (G_0W_0 proper retarded cumulant, or G_0W_0 PRC) [142], except that the improper self-energy is used in place of the proper self-energy. To calculate the Green's function within one of these two cumulant schemes, a proper or improper retarded self-energy is first computed as in the G_0W_0 scheme. Then the self-energy is inserted into the following Fourier- transformed version of Eq. (4.26) to find the cumulant:

$$\begin{aligned} C_k(t) &\equiv C_k(t_0, t_0+t) \\ &= \int d\omega \frac{1}{\pi} \frac{|\mathrm{Im}\Sigma_k^R(\omega+\epsilon_k)|}{\omega^2}(e^{-i\omega t} + i\omega t - 1). \end{aligned} \quad (4.27)$$

Finally, the spectral weight for the Green's function is obtained by taking a Fourier transform of Eq. (4.24) [142].

Unlike standard G_0W_0 and G_0W_0 PRC, the cumulant approach outlined above no longer sums

diagrams in order of the number of interactions, and is thus not perturbative in the interaction coupling. Instead, the first cumulant in Eq. 4.26 contains diagrams of all orders of the interaction. We emphasize that this fact renders the approach neither more or less accurate than the more standard G_0W_0 and G_0W_0 PRC approximations. Regardless, the simplicity of the cumulant formalism as outlined above does lead to important properties such as positive and normalized spectral weight [142].

4.4 Results

To gain a first understanding of the implications of the G_0W_0 IRC approximation, we apply it to the study of a system of independent phonons coupled to a single electronic state:

$$H = \sum_k \omega_k b_k^\dagger b_k + c^\dagger c \left[\epsilon_c + \sum_k M_k (b_k^\dagger + b_k) \right], \tag{4.28}$$

where $b_k, b_k^\dagger$ are the annihilation and creation operators for the phonon states, $c, c^\dagger$ are those of the electronic state, and ϵ_c, λ_k are the excited state electronic energy and the phonon coupling, respectively. We utilize an Einstein spectral density, $J(\omega) = g\delta(\omega - \omega_c)$. This is crudely reminiscent of the plasmon spectral density in the 3D electron gas. This model has been used previously previously in the analysis of interaction effects for core-holes [144] and to develop models for valence band spectral functions [126]. The model is solved exactly at zero temperature in the second Appendix to this chapter.

For this model Hamiltonian in which the interaction propagator W has been replaced by its phonon propagator analogue D, the G_0D_0 PRC approach gives the exact result for the spectral weight [21]. In this regard our goal is not to compare with the standard cumulant approach, which will of course in this model yield "better" results, but to gain an intuition for the expected spectral features produced by the IRC method as well as to see which features of the approach are likely to be well described.

The dynamical part of the proper Dyson self-energy within this model is

$$\begin{aligned}
\Sigma^*(\omega) &= \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega' g\omega_c^2 G^0(\omega-\omega';\epsilon_c)D^0(\omega';\omega_c), \\
&= \frac{i}{2\pi}\int_{-\infty}^{\infty} d\omega' g\omega_c^2 \frac{1}{\omega-\omega'-\epsilon_c+i\delta}\left(\frac{1}{\omega'-\omega_c+i\delta}-\frac{1}{\omega'+\omega_c-i\delta}\right), \\
&= \frac{g\omega_c^2}{\omega-\omega_c-\epsilon_c+i\delta}, \qquad (4.29) \\
|\text{Im }\Sigma^*(\omega)| &= \pi g\omega_c^2\delta(\omega-(\omega_c+\epsilon_c)). \qquad (4.30)
\end{aligned}$$

The frequency integration was done by closing the contour in the lower imaginary half-plane. The exact and approximate spectral functions are then evaluated as outlined in the previous section; the improper self-energy is described by

$$|\text{Im }\Sigma^I(\omega)| = \pi g(1+g)\omega_c^2\delta(\omega-\omega_c(1+g)-\epsilon_c), \qquad (4.31)$$

and the exact and approximate results are

$$\begin{aligned}
A_{\text{PRC}}(\omega) &= e^{-g}\sum_{l=0}^{\infty}\frac{g^l}{l!}\delta(\omega-\epsilon_c+g\omega_0-\omega_0 l), \qquad (4.32) \\
A_{G_0D_0}(\omega) &= \frac{1}{1+g}\delta(\omega-\epsilon_c+g\omega_0)+\frac{g}{1+g}\delta(\omega-\epsilon_c-\omega_0), \qquad (4.33) \\
A_{\text{IRC}}(\omega) &= e^{-g/(1+g)}\sum_{l=0}^{\infty}\frac{1}{l!}\left(\frac{g}{1+g}\right)^l\delta(\omega-\epsilon_c+g\omega_0-\omega_0 l(1+g)), \qquad (4.34)
\end{aligned}$$

respectively. The exact solution describes a sequence of peaks separated by multiples of ω_0 in energy. The basic G_0D_0 (Dyson) approximation predicts just two peaks, a quasiparticle peak and a satellite peak separated by $\omega_0(1+g)$, while the G_0D_0 IRC formulation predicts an infinite series of peaks separated by $\omega_0(1+g)$. The G_0D_0 IRC spectrum inherits the unphysical spacing from the G_0D_0 improper self-energy, which becomes correct only in the weak coupling ($g \ll 1$) limit.

It is easy to check that all three methods give normalized spectral weights with an identical first

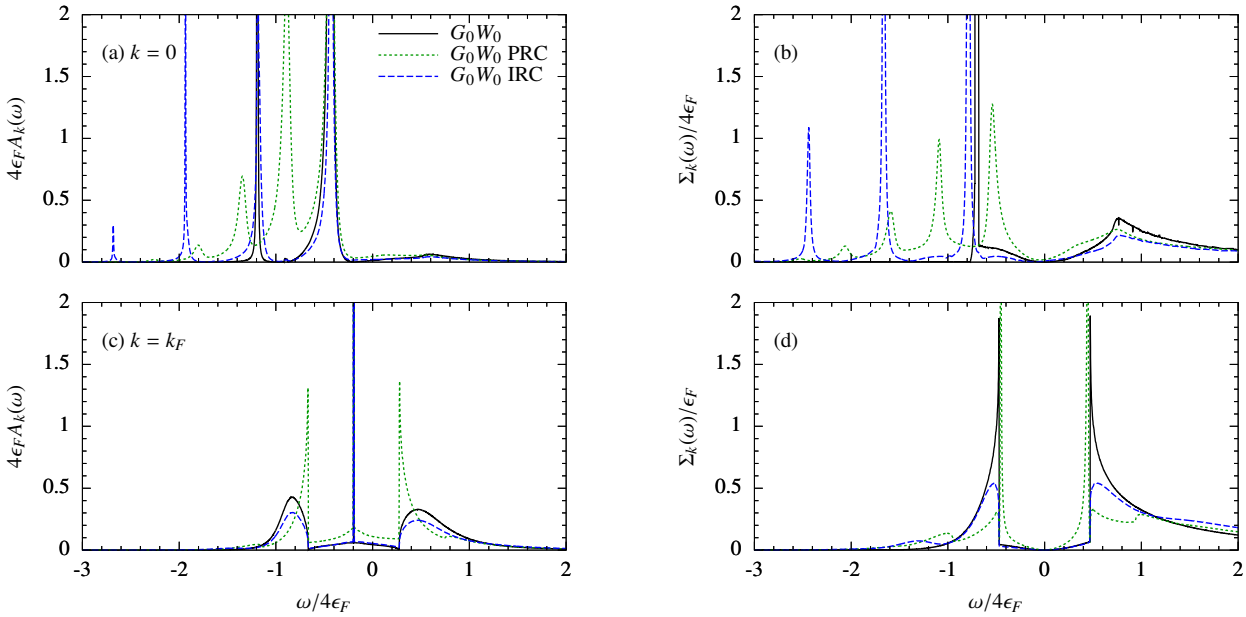

Figure 4.4: A comparison of three distinct GW schemes. (a,c) The spectral weight for the 3D electron gas with $r_s = 4.0$. (b,d) The absolute value of the imaginary part of the proper self-energy, solved using Dyson's equation, corresponding to the spectral weights in (a) and (c).

moment

$$\begin{aligned} \left\langle \left| cc^\dagger \left(\epsilon_c + \sum_k \lambda_k (b_k^\dagger + b_k) \right) \right| \right\rangle &= \int \omega A(\omega) d\omega \\ &= \epsilon_c, \end{aligned} \tag{4.35}$$

where $|\rangle$ represents the direct product of the ground electronic state and all ground phonon states. The G_0D_0 IRC spectrum, although inaccurate in its peak spacing, still encodes the correct first energy moment. Thus, the IRC approach appears, within this simple toy model, to embody a compromise between the standard cumulant and self-energy GW approaches. Furthermore, it appears not to corrupt some important aspects of the problem, such as the existence of multiple satellites and the value of the "correlation" energy.

The electron gas problem provides a more stringent and informative grounds for comparison of the different approximation schemes because no known method, including the G_0W_0 PRC scheme, yields the exact solution. Although the electron gas Hamiltonian differs significantly from the form of Eq. 4.28, many of the observations made earlier about the three approximations to the Green's

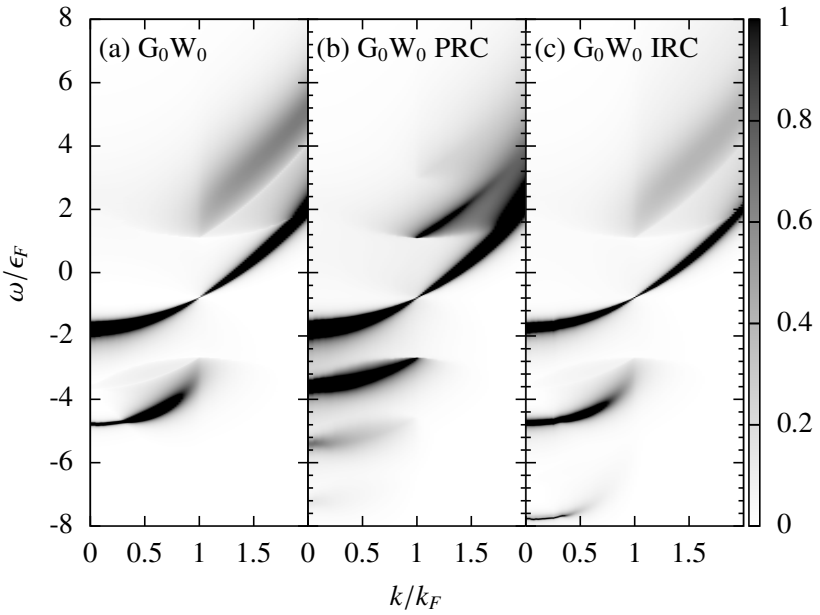

Figure 4.5: Band structure for various GW schemes at $r_s = 4$. Note that the G_0W_0 IRC contains multiple satellites, similar to the G_0W_0 PRC case, but with a spacing that mimics the incorrect position of the G_0W_0 plasmaron peak.

function will still apply. In Fig. 4.4a,c, we present the spectral weight

$$A_k(\omega) = -\frac{1}{\pi}\mathrm{Im}G_k^R(\omega) \tag{4.36}$$

for the 3D electron gas at density $r_s = 4.0$ both at the Fermi energy and at the bottom of the band. The spectral weight for the Dyson GW self-energy is plotted for the same cases in Fig. 4.4b,d. The

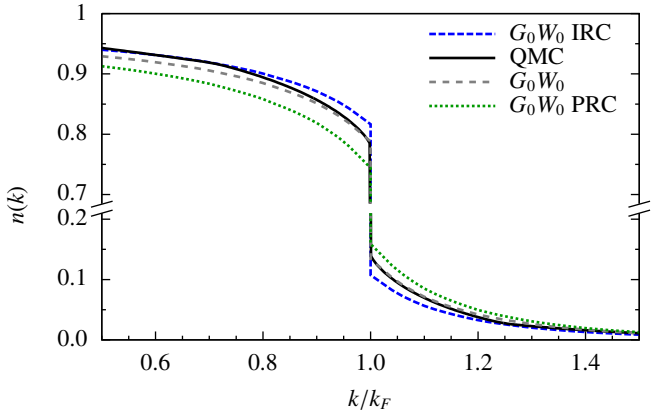

Figure 4.6: The $k-$dependent occupation number $n(k)$ for all three approximation schemes at $r_s = 4.0$, plotted against the "exact" QMC result [145, 146].

full band structure is reported for all three calculation schemes in Fig. 4.5.

Both cumulant schemes improve the qualitative description of the satellite region at low k: whereas the G_0W_0 scheme predicts just one satellite at an energy $1.5\omega_p$ below the quasiparticle (QP) peak, both cumulant schemes predict a series of evenly-spaced satellites with decreasing weights. The G_0W_0 PRC scheme predicts that these satellite peaks are separated by approximately the plasmon energy ω_p, whereas the G_0W_0 IRC scheme inherits the (presumably unphysical) $\sim$1.5ω_p spacing from the G_0W_0 calculation, exactly as it did in the electron-phonon model explored earlier. At and above the Fermi wavevector k_F, the G_0W_0 IRC results are much more similar to the G_0W_0 results as compared to those from the G_0W_0 PRC. Overall, the G_0W_0 IRC scheme interpolates between the rounded spectral behavior of G_0W_0 at larger k and the multiple satellite peak behavior of G_0W_0 PRC at lower k.

In addition to the spacing of satellite peaks, another unphysical feature of the IRC approach is the appearance of spurious sharp quasiparticle resonances which are most apparent in the unoccupied portion of the spectral function. These features can be easily removed in a manner that hardly affects $A_k(\omega)$ for $k \leq k_F$, $n(k)$, or ϵ_{corr}/N (the latter two of which are discussed nex). The origin of these features and the means for their removal are discussed in an Appendix of this chapter.

While the results for the 3D electron gas in the high frequency satellite wing suggest that the G_0W_0 PRC is most accurate in this spectral region, these results shed little light on other properties, to which we now turn. We present in Fig. 4.6 the wavevector-dependent occupation number

$$n_k = \int_{-\infty}^{\mu} A_k(\omega) d\omega. \tag{4.37}$$

The G_0W_0 IRC scheme performs similarly to G_0W_0, erring on the opposite side of the "exact" quantum Monte Carlo (QMC) data [145, 146]. Notably, the G_0W_0 IRC occupation numbers match the QMC results almost exactly for $k < 0.9k_F$ and $k > 1.1k_F$; although the scheme suffers from unphysical satellite peak spacing at small k, it inherits rather accurate occupation numbers from G_0W_0. The G_0W_0 PRC occupation numbers are not as accurate, and yield a QP renormalization

r_s	G_0W_0	G_0W_0 PRC	G_0W_0 IRC	QMC
1	-0.0742	-0.0688	-0.0642	-0.0600
2	-0.0542	-0.0516	-0.0467	-0.0448
3	-0.0436	-0.0411	-0.0368	-0.0369
4	-0.0375	-0.0350	-0.0310	-0.0318
5	-0.0329	-0.0304	-0.0267	-0.0281

Table 4.1: Correlation energies of the 3D electron gas per particle ϵ_{corr}/N in Hartrees for various GW schemes and r_s values [142]. QMC values are obtained from Vosco et al's parameterization [147] of Ceperley and Alder's fixed-node diffusion Monte Carlo data [128].

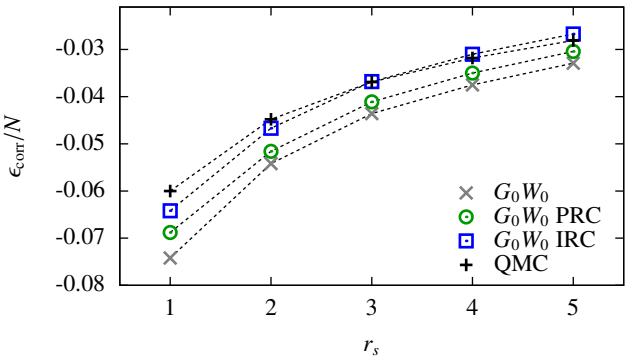

Figure 4.7: Correlation energies of the 3D electron gas per particle ϵ_{corr}/N for various GW schemes and r_s values compared with the "exact" QMC result. Exact values are reported in Table 4.1.

factor which is too small.

The accuracy of the occupation numbers in the G_0W_0 IRC approximation can be explained using the self-energy spectra in Fig. 4.4b,d. The most significant dependence of the momenta are on the weight of the main QP peak, which is determined by the slope of the real part of the self-energy. Since the real and imaginary parts of the self-energy are related by a Kramers-Kronig transform, the slope of the real part of the self-energy depends most strongly on the weight and positions of the peaks in the self-energy spectrum closest to $\omega \simeq \epsilon_k$. Since the G_0W_0 IRC self-energy spectrum peaks are more similar to the G_0W_0 ones, one would expect the G_0W_0 IRC momentum distribution near $k = k_F$ to more closely resemble that of G_0W_0. In particular, the increased QP renormalization factor with respect to the G_0W_0 approximation is explained by the slightly smaller peaks in the G_0W_0 IRC self-energy spectrum at $k = k_F$.

Finally, total energies may be calculated from $A_k(\omega)$ using the Galitskii-Migdal formula

$$\epsilon = \sum_k \int_{-\infty}^{\mu} (\omega + \epsilon_k) A_k(\omega) d\omega, \tag{4.38}$$

where $\epsilon_k = k^2/2$ is the free-electron energy dispersion. The correlation energy per particle is calculated using $\epsilon_{\text{corr}} = (\epsilon - \epsilon_{\text{HF}})/N$, where ϵ_{HF} is the Hartree-Fock energy. For the cumulant schemes, μ is determined by enforcing the total particle number. These energies are reported in Table I and plotted in Fig. 4.7. For reference, the results from fixed-node diffusion Monte Carlo calculations of Ceperley and Alder [128] are shown based on the parameterization by Vosko and coworkers [147]. The G_0W_0 IRC scheme yields significantly more accurate correlation energies as compared to earlier schemes over this important range of r_s values.

4.5 Conclusion

In this chapter, we motivate the use of the improper retarded self energy in the cumulant generating function using Dyson's equation. Non-self-consistent calculations of the spectral weight show that the improper retarded cumulant (IRC) scheme predicts a series of multiple satellite bands with a larger-than-expected spacing at the bottom of the band. However, compared to the G_0W_0 PRC scheme, which predicts a series of satellite bands with a more physical ω_p spacing, the IRC scheme yields noticeably improved occupation numbers and correlation energies. This is promising in the ongoing research directed to unified, efficient approaches for both total electronic energy and excitation energies. Further work should be done to investigate other aspects related to the retarded cumulant-based GW approaches discussed here, including self-consistency and the influence of higher-order cumulants.

4.6 Appendix: Spurious Sharp Quasiparticle Resonances at $k > k_F$

The standard description of the cumulant function requires the evaluation of

$$C_k(t) = \frac{1}{\pi}\int d\omega \frac{|\mathrm{Im}\Sigma(\omega+\epsilon_k)|}{\omega^2}(e^{-i\omega t} + i\omega t - 1), \tag{4.39}$$

which for the IRC demands that the integrand is constructed with $\mathrm{Im}\Sigma^{I,R}(\omega+\epsilon_k)$. It should be noted, however, that the geometric sequence that is the improper self-energy is ill-defined when evaluated precisely on the energy shell since $G_k^0(\epsilon_k)$ is divergent there. This divergence implies $\mathrm{Im}\Sigma^{I,R}(\epsilon_k) = 0$, resulting in a sharp quasiparticle-like feature in $A_k(\omega)$ superimposed on a smooth continuum. This feature is most apparent for $k > k_F$, as can be seen in Fig. 4.8a. Note that the smooth continuum behavior is much like that of the standard G_0W_0 spectral function.

This feature may be removed in a variety of ways that preserve all of the conclusions reached

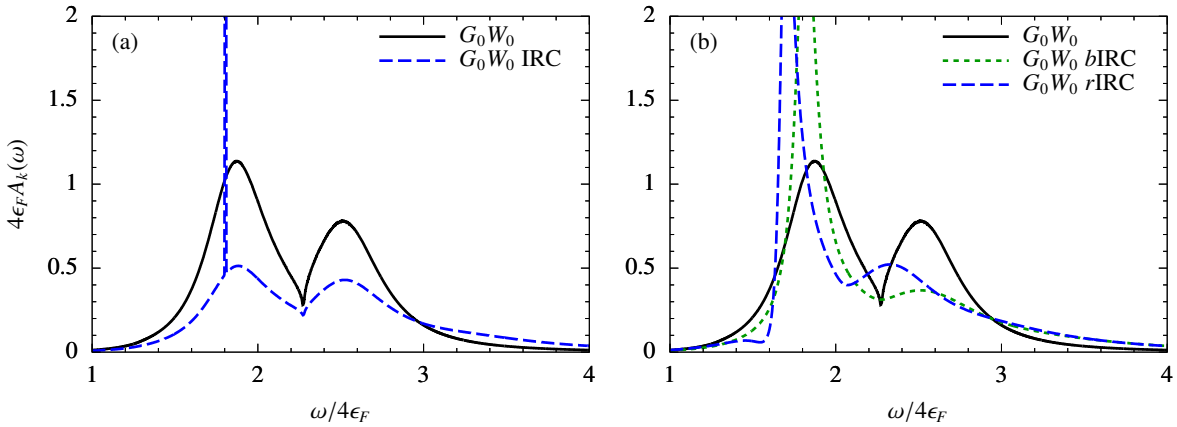

Figure 4.8: Band structure for various GW schemes at $r_s = 4$ and $k = 3k_F$. The sharp quasiparticle resonance predicted by G_0W_0 IRC is broadened by the bIRC and rIRC schemes.

r_s	G_0W_0 IRC	G_0W_0 bIRC	G_0W_0 rIRC	QMC
1	-0.0642	-0.0643		-0.0600
2	-0.0467	-0.0467		-0.0448
3	-0.0368	-0.0363		-0.0369
4	-0.0310	-0.0304	-0.0308	-0.0318
5	-0.0267	-0.0260		-0.0281

Table 4.2: Correlation energies of the 3D electron gas per particle $\epsilon_{\mathrm{corr}}/N$ in Hartrees for the broadened IRC (bIRC) and the renormalized IRC (rIRC) schemes.

in the main text. For example, if the cumulant function $C_k(t)$ is defined such that the evaluation of $\mathrm{Im}\Sigma_k(\omega + \epsilon_k^{G_0W_0})$ is used as opposed to $\mathrm{Im}\Sigma_k(\omega + \epsilon_k)$, then no spurious resonance appears. In addition, $A_k(\omega)$ for $k \leq k_F$, $n(k)$, and $\epsilon_{\mathrm{corr}}/N$ are essentially unchanged. The same outcome occurs if the frequency of $G_k^0(\omega)$ within $\Sigma_k^{I,R}(\omega)$ is given an imaginary part equal to $\Sigma_k^*(\epsilon_k^{G_0W_0})$, the quasiparticle lifetime associated with a standard G_0W_0 calculation.

These distinct regularization procedures do however alter the shape of $A_k(\omega)$ for $k > k_F$ as shown in Fig. 4.8b. Table 4.2 shows the robustness of the correlation energies for the different schemes; we do not replot $A_k(\omega)$ for $k \leq k_F$ or $n(k)$ because they are essentially unaltered from Figs. 4.4,4.5,4.6. Overall, we find that the most important features of the IRC approach are basically unmodified under the removal of these spurious resonances.

4.7 Appendix: Derivation of Electron-Phonon Green's Function

In this section, we derive the exact zero-temperature solution for the electron-phonon model considered in this work. The solution for the $T > 0$ case contains mostly the same ingredients and can be found in [21].

The model can be written as

$$\begin{aligned} H &= \sum_k \omega_k b_k^\dagger b_k + c^\dagger c \left[\epsilon_c + \sum_k M_k(b_k^\dagger + b_k)\right] \\ &= \Big|g\Big\rangle \sum_k \omega_k b_k^\dagger b_k \Big\langle g\Big| + \Big|e\Big\rangle \left[\epsilon_c + \sum_k (\omega_k b_k^\dagger b_k + M_k(b_k^\dagger + b_k))\right] \Big\langle e\Big| \\ &\equiv |g\rangle H_g \langle g| + |e\rangle H_e \langle e|, \end{aligned} \tag{4.40}$$

where $|g\rangle$ refers to the ground electronic state and $|e\rangle$ refers to the excited electronic state (i.e. $c^\dagger|g\rangle = |e\rangle$ and $c|e\rangle = |g\rangle$). Because $(b^\dagger + b) \propto x$, the position operator, the excitation of the electron has the effect of creating an electric field that couples to each phonon mode.

The Green's function is defined as

$$\begin{aligned} G(t) &= -i\langle g;\{\}|c(t)c^\dagger(0)|g;\{\}\rangle \\ &= -i\langle g;\{\}|e^{iHt/\hbar}ce^{-iHt/\hbar}c^\dagger|g;\{\}\rangle. \end{aligned} \tag{4.41}$$

where $|g,\{\}\rangle$ refers to the direct product of the ground electronic state with all phonon states, each having zero excitations.

Applying all operators to the right in turn, the Green's function can be written simply as

$$G(t) = -i\langle\{\}|e^{iH_gt/\hbar}e^{-iH_et/\hbar}|\{\}\rangle. \tag{4.42}$$

The electronic index has been dropped in the wavefunctions because the operators H_g, H_e can be considered to act solely on the phonon states. We can immediately make two simplifications:

$$G(t) = -ie^{-i\epsilon_ct/\hbar}\prod_k \left\langle n_k = 0\middle|e^{i\omega_kb_k^\dagger b_kt/\hbar}e^{-i(\omega_kb_k^\dagger b_k+M_k(b_k^\dagger+b_k))t/\hbar}\middle|n_k=0\right\rangle. \tag{4.43}$$

We've factored out the phase $e^{-i\epsilon_ct/\hbar}$ due to the electronic excitation as well as split the Green's function into a product of contributions from each phonon mode. The latter is possible because the phonon creation and annihilation operators for different momenta commute with one another.

For simplicity, we henceforth consider only a single term of the product in Eq. 4.43. Since $|n_k = 0\rangle$ is an eigenstate of the single-phonon hamiltonian $\omega_kb_k^\dagger b_k$, the first exponential can be applied to the left to give 1, i.e.

$$\langle 0|e^{i\omega_kb_k^\dagger b_kt/\hbar}e^{-i(\omega_kb_k^\dagger b_k+M_k(b_k^\dagger+b_k))t/\hbar}|0\rangle = \langle 0|e^{-i(\omega_kb_k^\dagger b_k+M_k(b_k^\dagger+b_k))t/\hbar}|0\rangle. \tag{4.44}$$

We can rewrite the remaining exponential by first completing the square:

$$\begin{aligned}\omega_k b_k^\dagger b_k + M_k(b_k^\dagger + b_k) &= \omega_k\left(b_k^\dagger b_k + \frac{M_k}{\omega_k}(b_k^\dagger + b_k)\right) \\ &= \omega_k\left(b_k^\dagger + \frac{M_k}{\omega_k}\right)\left(b_k + \frac{M_k}{\omega_k}\right) - \frac{M_k^2}{\omega_k}.\end{aligned} \tag{4.45}$$

The latter constant contributes the term $\exp(i(M_k^2/\omega_k)t/\hbar)$ to the Green's function; the former portion remains to be evaluated between the $|0\rangle$ bra and ket. Write $M_k/\omega_k \equiv \lambda_k$.

To begin, we show

$$\exp(-i\omega_k(b_k^\dagger + \lambda_k)(b_k + \lambda_k)t/\hbar) = e^{-S_k}\exp(-i\omega_k b_k^\dagger b_k t/\hbar)e^{S_k}, \tag{4.46}$$

where $S_k = \lambda_k(b_k^\dagger - b_k)$. Since $S_k^\dagger = -S_k$, we can insert the identity operator $1 = e^{S_k}e^{-S_k}$ between each pair of operators in the power series expansion for $e^{\omega_k b_k^\dagger b_k}$. In this way, we see $\overline{\exp(-i\omega_k b_k^\dagger b_k t/\hbar)} = \exp(-i\omega_k \overline{b_k^\dagger}\,\overline{b_k} t/\hbar)$, where $\overline{A}$ denotes $e^{-S_k}Ae^{S_k}$.

We evaluate $\overline{b_k}$ using direct applications of the commutator relation $[b_k, b_k^\dagger] = 1$, which implies $[b_k, S_k] = \lambda_k$.

$$\begin{aligned} b_k e^{S_k} &= b_k\left(1 + S_k + \frac{S_k^2}{2!} + \frac{S_k^3}{3!} + \ldots\right) \\ &= b_k + (S_k b_k + \lambda_k) + \frac{1}{2!}(S_k b_k + \lambda_k)S_k + \frac{1}{3!}(S_k b_k + \lambda_k)S_k^2 + \ldots \\ &= b_k + (S_k b_k + \lambda_k) + \frac{1}{2!}(2\lambda_k S_k + S_k^2 b_k) + \frac{1}{3!}(3\lambda_k S_k^2 + S_k^3 b_k) + \ldots \\ &= e^{S_k}(b_k + \lambda_k). \end{aligned} \tag{4.47}$$

Similarly, $\overline{b_k^\dagger} = b_k^\dagger + \lambda_k$ since $[b_k^\dagger, S_k] = \lambda_k$ as well, so that Eq. 4.46 has been proven.

The task has been reduced to evaluating

$$\left\langle 0\left| e^{-\lambda_k(b_k^\dagger - b_k)} e^{-i\omega_k b_k^\dagger b_k t/\hbar} e^{\lambda_k(b_k^\dagger - b_k)}\right| 0\right\rangle. \tag{4.48}$$

To proceed, we apply a simple case of the Baker-Campbell-Hausdorff formula which states that

for operators A, B satisfying $[A,[A,B]] = [B,[A,B]] = 0$, the relation $e^{A+B} = e^A e^B e^{-[A,B]/2}$ holds.

Applying the BCH formula to the canonical transformation operators in Eq. 4.48 gives

$$e^{\lambda_k (b_k^\dagger - b_k)} = e^{\lambda_k b_k^\dagger} e^{-\lambda_k b_k} e^{-\lambda_k^2/2}, \tag{4.49}$$

so that the expression in Eq. 4.48 can be written as

$$e^{-\lambda_k^2} \left\langle 0 \middle| e^{-\lambda_k b_k^\dagger} e^{\lambda_k b_k} e^{-i\omega_k b_k^\dagger b_k t/\hbar} e^{\lambda_k b_k^\dagger} e^{-\lambda_k b_k} \middle| 0 \right\rangle. \tag{4.50}$$

The two exponential operators towards the outside act as one because any higher term in their power series will annihilate the bra or ket. Thus, we are left with

$$\begin{aligned}
\left\langle 0 \middle| e^{\lambda_k b_k} e^{-i\omega_k b_k^\dagger b_k t/\hbar} e^{\lambda_k b_k^\dagger} \middle| 0 \right\rangle &= \left\langle 0 \middle| \left(\sum_n \frac{(\lambda_k b_k)^n}{n!} \right) e^{-i\omega_k b_k^\dagger b_k t/\hbar} \left(\sum_n \frac{(\lambda_k b_k^\dagger)^n}{n!} \right) \middle| 0 \right\rangle \\
&= \left(\sum_n \frac{\lambda_k^n}{\sqrt{n!}} \langle n| \right) e^{-i\omega_k b_k^\dagger b_k t/\hbar} \left(\sum_n \frac{\lambda_k^n}{\sqrt{n!}} |n\rangle \right) \\
&= \sum_n \frac{\lambda_k^{2n}}{n!} e^{-in\omega_k t/\hbar} \langle n|n\rangle \\
&= \exp\left(\lambda_k^2 e^{-i\omega_k t/\hbar} \right).
\end{aligned} \tag{4.51}$$

Collecting all of the factors that we have accumulated, we see

$$G(t) = -ie^{-i\epsilon_c t/\hbar} \prod_k \exp\left(\frac{M_k^2}{\omega_k^2} \left(i\omega_k t/\hbar - 1 + e^{-i\omega_k t/\hbar} \right) \right). \tag{4.52}$$

Now, within the Einstein model, all phonon frequencies are taken to be equal, i.e. $\omega_k = \omega_0$. In this case, we can simplify Eq. 4.52 by introducing the coupling constant

$$g = \sum_k \frac{M_k^2}{\omega_k^2}. \tag{4.53}$$

Then the Green's function of time simplifies to

$$G(t) = -ie^{-i(\epsilon_c - g\omega_0)t/\hbar} e^{-g(1-e^{-i\omega_0 t/\hbar})}. \tag{4.54}$$

To compute the Green's function as a function of frequency, we expand the double exponential in a power series and take a Fourier transform:

$$\begin{aligned} G(\omega) &= -ie^{-g} \int dt e^{-i(\epsilon_c - g\omega_0 - \omega)t/\hbar} \sum_n \frac{g^n}{n!} e^{-in\omega_0 t/\hbar} && (4.55) \\ &= e^{-g} \sum_n \frac{g^n}{n!} \delta(\omega - (\epsilon_c - g\omega_0 + n\omega_0)), && (4.56) \end{aligned}$$

which matches Eq. 4.32 exactly.

Physically, before the electron excitation is created, the phonons are in their ground states. At the moment the excitation is created, the "equilibrium point" of each of the oscillators becomes shifted so that the system wavefunction is now a superposition of new phonon eigenstates. Each of these parts of the superposition is propagated in according to its energy, which is why the energies of delta functions in the Green's function are spaced by the amount $g\omega_0$. The amplitude of each delta function corresponds to the sum of the overlap between the initial system wavefunction and the new boson eigenstates of the same energy.

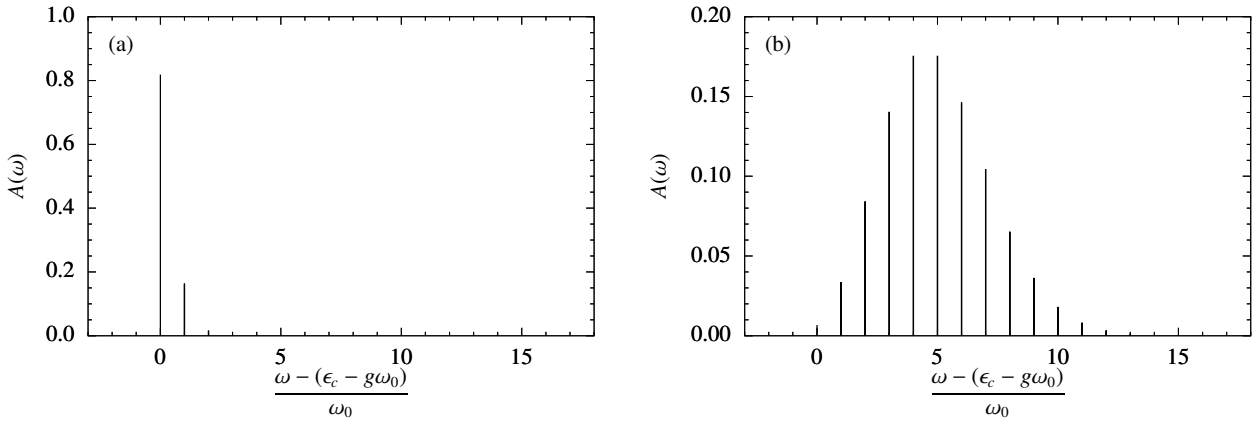

Figure 4.9: Spectral weight for the Green's function of the electron-phonon model within the Einstein phonon model for (a) $g = 0.2$ and (b) $g = 5$. Each vertical line represents a Dirac delta function with weight corresponding to the height of the line.

In Fig. 4.9, we plot spectral weights for two different values of the coupling constant g. For low coupling constants (panel a), the system evolves mostly in the ground eigenstate of the new phonon hamiltonian because the positions of the oscillators have not shifted a great amount ($g < 1$). For larger coupling constants, the oscillators are shifted enough that the excited phonon eigenstates play a more important role.

BIBLIOGRAPHY

1. J.-P. Correa-Baena *et al.*, *Science* **358**, 739–744 (2017).

2. M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami, H. J. Snaith, *Science* **338**, 643–647 (2012).

3. M. Liu, M. B. Johnston, H. J. Snaith, *Nature* **501**, 395 EP – (Sept. 2013).

4. J. Burschka *et al.*, *Nature* **499**, 316 EP – (July 2013).

5. F. Deschler *et al.*, *J. Phys. Chem. Lett.* **5**, 1421–1426 (Apr. 2014).

6. Z.-K. Tan *et al.*, *Nat. Nano.* **9**, 687–692 (2014).

7. B. R. Sutherland, E. H. Sargent, *Nature Photonics* **10**, 295 EP – (Apr. 2016).

8. G. Xing *et al.*, *Nature Materials* **13**, 476 EP – (Mar. 2014).

9. S. D. Stranks, H. J. Snaith, *Nature Nanotechnology* **10**, 391 EP – (May 2015).

10. P. S. Whitfield *et al.*, *Scientific Reports* **6**, 35685 EP – (Oct. 2016).

11. T. Baikie *et al.*, *J. Mater. Chem. A* **3**, 9298–9307 (17 2015).

12. A. N. Beecher *et al.*, *ACS Energy Letters* **1**, 880–887 (2016).

13. O. Yaffe *et al.*, *Phys. Rev. Lett.* **118**, 136001–1–6 (2017).

14. L. D. Whalley, J. M. Skelton, J. M. Frost, A. Walsh, *Phys. Rev. B* **94**, 220301 (Dec. 2016).

15. M. Sendner *et al.*, *Mater. Horiz.* **3**, 613–620 (6 2016).

16. X. Wu *et al.*, *Science Advances* **3**, e1602388 (2017).

17. M. Zhang, X. Zhang, L.-Y. Huang, H.-Q. Lin, G. Lu, *Phys. Rev. B* **96**, 195203 (19 2017).

18. J. M. Frost, *Phys. Rev. B* **96**, 195202 (19 2017).

19. S. De Wolf *et al.*, *Journal of Physical Chemistry Letters* **5**, 1035–1039 (Mar. 2014).

20. F. Zheng, H. Takenaka, F. Wang, N. Z. Koocher, A. M. Rappe, *The Journal of Physical Chemistry Letters* **6**, PMID: 26263087, 31–37 (2015).

21. G. D. Mahan, *Many-Particle Physics* (Kluwer Academic/Plenum Publishers, 3rd, 2000).

22. J. Even, L. Pedesseau, J.-M. Jancu, C. Katan, *J. Phys. Chem. Lett.* **4**, 2999–3005 (2013).

23. F. Zheng, L. Z. Tan, S. Liu, A. M. Rappe, *Nano Lett.* **15**, PMID: 26461166, 7794–7800 (2015).

24. P. Giannozzi *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).

25. A. M. Rappe, K. M. Rabe, E. Kaxiras, J. D. Joannopoulos, *Phys. Rev. B Rapid Comm.* **41**, 1227–1230 (1990).

26. N. J. Ramer, A. M. Rappe, *Phys. Rev. B* **59**, 12471–12478 (1999).

27. A. A. Mostofi *et al.*, *Comput. Phys. Commun.* **178**, 685 (2008).

28. N. Marzari, D. Vanderbilt, *Phys. Rev. B.* **56**, 12847–12865 (1997).

29. A Mattoni, A Filippetti, M. I. Saba, P Delugas, *J. Phys. Chem. C* **119**, 17421–17428 (2015).

30. W. A. Saidi, S. Poncé, B. Monserrat, *J. Phys. Chem. Lett.* **7**, PMID: 27973908, 5247–5252 (2016).

31. J. A. e. Guerra, *J. Phys. Chem. Lett.* **7**, 173104 (2017).

32. S. P. Senanayak *et al.*, *Sci. Adv.* **3**, eprint: `http://advances.sciencemag.org/content/3/1/e1601935.full.pdf` (2017).

33. H. Zhu, V. May, in *Energy Transfer Dynamics in Biomaterial Systems*, ed. by I. Burghardt, V. May, D. A. Micha, E. R. Bittner (Springer Series in Chemical Physics, 2009), vol. 93, p. 35.

34. L. Wang, D. Beljonne, L. Chen, Q. Shi, *J. Chem. Phys.* **134**, 244116 (2011).

35. S. Shrestha *et al.*, *The Journal of Physical Chemistry C* **122**, 5935–5939 (2018).

36. T. J. Savenije *et al.*, *The Journal of Physical Chemistry Letters* **5**, PMID: 26279532, 2189–2194 (2014).

37. M. Karakus *et al.*, *The Journal of Physical Chemistry Letters* **6**, PMID: 26619006, 4991–4996 (2015).

38. Y. H. Taek, W. Xiaoxi, Z. Xiaoyang, P. Vitaly, *Advanced Materials* **28**, 6509–6514 (2016).

39. J. S. Bader, R. A. Kuharski, D. Chandler, *J. Chem. Phys.* **93**, 230–236 (1990).

40. K. Binder, *Phys. Rev. Lett.* **47**, 693–696 (9 1981).

41. A. Splendiani *et al.*, *Nano Lett.* **10**, 1271–1275 (2010).

42. K. Mak, C. Lee, J. Hone, J. Shan, T. F. Heinz, *Phys. Rev. Lett.* **105**, 136805 (2010).

43. S. Z. Butler *et al.*, *ACS Nano.* **7**, 2898–2926 (2013).

44. A. K. Geim, I. V. Grigorieva, *Nature* **499**, 419–425 (2013).

45. M. Chhowalla *et al.*, *Nat. Chem.* **5**, 263–275 (2013).

46. D. Jariwala, V. K. Sangwan, L. J. Lauhon, T. J. Marks, M. C. Hersam, *ACS Nano.* **8**, 1102–1120 (2014).

47. T. S. Li, G. L. Galli, *J. Phys. Chem. C* **111**, 16192–16196 (2007).

48. S. Lebegue, O. Eriksson, *Phys. Rev. B* **79**, 115409 (2009).

49. O. Gunawan *et al.*, *Phys. Rev. Lett.* **97**, 186404 (2006).

50. K. F. Mak, K. L. He, J. Shan, T. F. Heinz, *Nat. Nano.* **7**, 494–498 (2012).

51. H. L. Zeng, J. F. Dai, W. Yao, D. Xiao, X. D. Cui, *Nat. Nano.* **7**, 490–493 (2012).

52. D. Xiao, G. B. Liu, W. X. Feng, X. D. Xu, W. Yao, *Phys. Rev. Lett.* **108**, 196802 (2012).

53. T. Cao, G. Wang, W. P. Han, H. Q. Ye, *Nat. Commun.* **3**, 887 (2012).

54. A. Chernikov *et al.*, *Phys. Rev. Lett.* **113**, 076802 (2014).

55. K. L He *et al.*, *Phys. Rev. Lett.* **113**, 026803 (2014).

56. M. M. Ugeda *et al.*, *Nat. Mater.* **13**, 1091–1095 (2014).

57. Y. You *et al.*, *Nat. Phys.* **11**, 477–481 (2015).

58. K. F. Mak *et al.*, *Nat. Mat.* **12**, 207–211 (2013).

59. J. Z. Shang *et al.*, *ACS Nano* **9**, 647–655 (2015).

60. J. S. Ross *et al.*, *Nat. Commun.* **4**, 1474 (2013).

61. M. S. Hybertsen, S. G. Louie, *Phys. Rev. B* **34**, 5390–5413 (1986).

62. M. Rohlfing, S. G. Louie, *Phys. Rev. B* **62**, 4927–4944 (2000).

63. T. Cheiwchanchamnangij, W. R. L. Lambrecht, *Phys. Rev. B* **85**, 205302 (2012).

64. A. Ramasubramaniam, *Phys. Rev. B* **86**, 115409 (2012).

65. D. Y. Qiu, F. H. da Jornada, S. G. Louie, *Phys. Rev. Lett.* **111**, 216805 (2013).

66. W. Hanke, L. J. Sham, *Phys. Rev. B* **21**, 4656–4673 (1980).

67. L. V. Keldysh, *JETP Lett.* **29**, 658–661 (1979).

68. P. Cudazzo, I. V. Tokatly, A. Rubio, *Phys. Rev. B* **84**, 085406 (2011).

69. P. Cudazzo, C. Attaccalite, I. V. Tokatly, A. Rubio, *Phys. Rev. Lett.* **104**, 226804 (2010).

70. T. C. Berkelbach, M. S. Hybertsen, D. R. Reichman, *Phys. Rev. B* **88**, 045318 (2013).

71. G. Berghauser, E. Malic, *Phys. Rev. B* **89**, 125309 (2014).

72. C. J. Zhang, H. N. Wang, W. M. Chan, C. Manolatou, F. Rana, *Phys. Rev. B* **89**, 205436 (2014).

73. M. M. Fogler, L. V. Butov, K. S. Novoselov, *Nat. Commun.* **5**, 4555 (2014).

74. L. Wang, A. Kutana, B. I. Yakobson, *Annalen Der Physik* **526**, L7–L12 (2014).

75. F. Wu, F. Qu, A. H. MacDonald, *Phys. Rev. B* **91**, 075310 (2015).

76. F. Huser, T. Olsen, K. S. Thygesen, *Phys. Rev. B* **88**, 245309 (2013).

77. A. Steinhoff, M. Rosner, F. Jahnke, T. O. Wehling, C. Gles, *Nano Lett.* **14**, 3743–3748 (2014).

78. T. C. Berkelbach, M. S. Hybertsen, D. R. Reichman, *Phys. Rev. B (in press)* **arXiv:1505.07127** (2015).

79. B. Ganchev, N. Drummond, I. Aliener, V. Fal'ko, *Phys. Rev. Lett.* **114**, 107401 (2015).

80. W. M. C. Foulkes, L. Mitas, R. J. Needs, G. Rajagopal, *Rev. Mod. Phys.* **73**, 33–83 (2001).

81. R. C. Grimm, R. G. Storer, *J. Comp. Phys.* **7**, 134 (1971).

82. M. H. Kalos, D. Levesque, L. Verlet, *Phys. Rev. A* **9**, 2178–2195 (1974).

83. C. J. Umrigar, K. G. Wilson, J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1719–1722 (1988).

84. N. D. Drummond, R. J. Needs, *Phys. Rev. B* **72**, 085124 (2005).

85. H. Flyvbjerg, H. G. Petersen, *J. Chem. Phys.* **91**, 461 (1989).

86. Y. Zhang *et al.*, *ACS Nano.* **Article ASAP**, 10.1021/acsnano.5b03505 (2015).

87. J. S. Ross *et al.*, *Nat. Commun.* **4** (2013).

88. G. Plechinger *et al.*, *arXiv:1507.01342* (2015).

89. B. Zhu, H. Zeng, J. Dai, Z. Gong, X. Cui, *Proc. Natl. Acad. Sci.* **111**, 11606–11611 (2014).

90. A. M. Jones *et al.*, *Nat. Nano.* **8**, 634–638 (2013).

91. C. Mai *et al.*, *Nano. Lett.* **14**, 202–206 (2014).

92. R. N. Barnett, P. J. Reynolds, W. A. Lester, *J. Comp. Phys.* **96**, 258–276 (1991).

93. W. F. Brinkman, T. M. Rice, B. Bell, *Phys. Rev. B* **8**, 1570 (1973).

94. D. A. Kleinman, *Phys. Rev. B* **28**, 871 (1983).

95. D. Birkedal, J. Singh, V. G. Lyssenko, J. Erland, J. M. Hvam, *Phys. Rev. Lett.* **76**, 672 (1996).

96. S. Schmitt-Rink, D. S. Chemla, D. A. B. Miller, *Adv. Phys.* **38**, 89 (1989).

97. K. Velizhanin, A. Saxena, *arXiv:1505.03910v1* (2015).

98. S. Chandrasekhar, *Astrophys. J.* **100**, 176 (1944).

99. M. Barbone *et al.*, *arXiv:1805.04950* (2015).

100. O. A. Ajayi *et al.*, *2D Materials* **4**, 031011 (2017).

101. G. Ramon, A. Mann, E. Cohen, *Phys. Rev. B* **67**, 045323 (4 2003).

102. S. Chandrasekhar, *Astrophys. J.* **100**, 176 (1944).

103. A. Chernikov *et al.* (*(to be submitted)*).

104. Y. W. Wen, A. Chernikov, M. Mayers, M. Glazov, A. Raja (*(to be submitted)*).

105. A. J. Cohen, P. Mori-Sánchez, W. Yang, *Chem. Rev.* **112**, 289–320 (2012).

106. G. Onida, L. Reining, A. Rubio, *Mod. Rev. Phys.* **74**, 601–659 (2002).

107. R. J. Bartlett, M. Musiał, *Rev. Mod. Phys.* **79**, 291–352 (2007).

108. G. H. Booth, A. Gruneis, G. Kresse, A. Alavi, *Nature* **493**, 365–370 (2013).

109. J. J. Shepherd, A. Grüneis, *Phys. Rev. Lett.* **110**, 226401 (2013).

110. J. McClain *et al.*, *arXiv:1512.04556* (2015).

111. L. Hedin, S. Lundqvist, *Solid State Phys.* **23**, 1 (1969).

112. L. Hedin, *Phys. Rev.* **139**, A796–A823 (1965).

113. M. S. Hybertsen, S. G. Louie, *Phys. Rev. B* **34**, 5390–5413 (1986).

114. R. W. Godby, M. Schlüter, L. J. Sham, *Phys. Rev. B* **35**, 4170 (1987).

115. F. Aryasetiawan, O. Gunnarsson, *Rep. Prog. Phys.* **61**, 237–312 (1998).

116. W. G. Aulbur, L. Jönsson, J. W. Wilkins, *Solid State Phys.* **54**, ed. by H. Ehrenreich, F. Spaepen, 1 –218 (1999).

117. E. L. Shirley, *Phys. Rev. B* **54**, 7758–7764 (1996).

118. U. von Barth, B. Holm, *Phys. Rev. B* **54**, 8411–8419 (1996).

119. B. Holm, U. von Barth, *Phys. Rev. B* **57**, 2108–2117 (1998).

120. F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, L. Reining, *Phys. Rev. Lett.* **94**, 186402 (2005).

121. M. van Schilfgaarde, T. Kotani, S. Faleev, *Phys. Rev. Lett.* **96**, 226402–4 (2006).

122. M. Shishkin, M. Marsman, G. Kresse, *Phys. Rev. Lett.* **99**, 246403–4 (2007).

123. M. Guzzo *et al.*, *Phys. Rev. Lett.* **107**, 166401 (2011).

124. A. Grüneis, G. Kresse, Y. Hinuma, F. Oba, *Phys. Rev. Lett.* **112**, PRL, 096401 (2014).

125. K. Van Houcke, I. S. Tupitsyn, A. S. Mishchenko, N. V. Prokof'ev, *Phys. Rev. B* **95**, 195131 (19 2017).

126. B. Holm, F. Aryasetiawan, *Phys. Rev. B* **62**, 4858–4865 (2000).

127. N. E. Dahlen, R. van Leeuwen, U. von Barth, *Phys. Rev. A* **73**, 012511–13 (2006).

128. D. M. Ceperley, B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).

129. B. I. Lundqvist, *Phys. Kondens. Materie* **7**, 117–123 (1968).

130. R. Kubo, *J. Phys. Soc. Jpn.* **17**, 1100–1120 (1962).

131. L. Hedin, *Phys. Scr.* **21**, 477–480 (1980).

132. C.-O. Almbladh, L. Hedin, in, ed. by E. E. Koch (North-Holland, 1983), vol. 1b, pp. 607–904, ISBN: 0444867104.

133. F. Aryasetiawan, L. Hedin, K. Karlsson, *Phys. Rev. Lett.* **77**, 2268–2271 (1996).

134. B. Holm, F. Aryasetiawan, *Phys. Rev. B* **56**, 12825 (1997).

135. M. Vos *et al.*, *Phys. Rev. B* **66**, 155414 (2002).

136. J. Lischner, D. Virgil-Fowler, S. G. Louie, *Phys. Rev. B* **89**, 125430 (2014).

137. J. Lischner, D. Vigil-Fowler, S. G. Louie, *Phys. Rev. Lett.* **110**, 146801 (2013).

138. M. Guzzo *et al.*, *Phys. Rev. B* **89**, 085425 (2014).

139. F. Caruso, H. Lambert, F. Guistino, *Phys. Rev. Lett.* **114**, 146404 (2015).

140. F. Caruso, F. Guistino, *Phys. Rev. B* **92**, 045123 (2015).

141. J. S. Zhou *et al.*, *JCP* **143**, 184019 (2015).

142. J. J. Kas, J. J. Rehr, L. Reining, *Phys. Rev. B* **90**, 085112 (2014).

143. A. L. Fetter, J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Dover Publications, Inc., 2003).

144. D. C. Langreth, *Phys. Rev. B* **1**, 471–477 (1970).

145. M. Holzmann *et al.*, *Phys. Rev. Lett.* **107**, 110402 (2011).

146. V. Olevano *et al.*, *Phys. Rev. B* **86**, 195123 (2012).

147. S. H. Vosko, L. Wilk, M. Nusair, *Canadian Journal of Physics* **58**, 1200–1211 (1980).